SYM-AM-17-035

PROCEEDINGS
OF THE
FOURTEENTH ANNUAL ACQUISITION RESEARCH SYMPOSIUM

THURSDAY SESSIONS
VOLUME II

Acquisition Research:
Creating Synergy for Informed Change

April 26–27, 2017

Published March 31, 2017

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

The research presented in this report was supported by the Acquisition Research Program of the Graduate School of Business & Public Policy at the Naval Postgraduate School.

To request defense acquisition research, to become a research sponsor, or to print additional copies of reports, please contact any of the staff listed on the Acquisition Research Program website (www.acquisitionresearch.net).

Table of Contents

SYM-AM-17-035

PROCEEDINGS
OF THE
FOURTEENTH ANNUAL ACQUISITION RESEARCH SYMPOSIUM

THURSDAY SESSIONS
VOLUME II

Acquisition Research:
Creating Synergy for Informed Change

April 26–27, 2017

Published March 31, 2017

THIS PAGE INTENTIONALLY LEFT BLANK

Keynote: Dr. Richard Carlin, Head, Sea Warfare and Weapons Department, Office of Naval Research

Richard T. Carlin—is Department Head of the Sea Warfare and Weapons Department at the Office of Naval Research (ONR). As department head, he oversees a broad range of science and technology (S&T) programs for surface ships, submarines, and undersea weapons with an annual budget of approximately $500 million per year.

Carlin entered the Senior Executive Service in January 2002 and has 14 years of federal service.

Prior to his current position as department head, Carlin was the director for the Undersea Weapons and Naval Materials Division with responsibilities in undersea weapons and countermeasures, advanced energetics, structural materials, materials for power systems, acoustic transducers, maintenance reduction technologies, and blast mitigation materials. During his career at ONR, he also served as the acting chief scientist in 2004 and as director for the Mechanics and Energy Conversion Division from 2001 to 2005. Prior to his appointment as a division director, Carlin was the program officer for Electrochemistry S&T and Undersea Weapons Propulsion with programs covering numerous electrochemical and thermal power technologies.

Additionally, Carlin serves as the Navy S&T executive on numerous Navy, Department of Defense, and interagency energy advisory groups, including the Navy's Task Force Energy Executive Steering Committee, DDR&E's Energy Security Task Force, and the Hydrogen and Fuel Cell Interagency Task Force. He also serves as a U.S. panel member on the NATO RTO Applied Vehicle Technology Panel, and is a member of the Department of the Navy Awards Review Panel.

Before joining ONR in August 1997, Carlin held several positions in academia, industry, and government. From 1995 to 1997, he was a senior scientist at Covalent Associates, Inc., performing contract research in areas of lithium batteries, supercapacitors, and ionic liquids catalysis. From 1992 to 1995, Carlin held the position of Electrochemistry Division chief at the Frank J. Seiler Research Laboratory (FJSRL) located at the U.S. Air Force Academy in Colorado Springs, CO. At FJSRL, he led research on the use of ionic liquids as electrolytes for batteries, supercapacitors, and metal-alloy electrodepositions, and as solvents for gas absorption and catalysis.

Carlin was an assistant professor of chemistry at the University of Alabama in Tuscaloosa from 1989 to 1992 where he taught both undergraduate and graduate level courses, and directed a research program in the study and application of ionic liquids as solvents and electrolytes. From 1982 to 1985, he was employed at Air Products and Chemicals as a senior research chemist carrying out research on the use of ionic liquids as gas-separation membranes.

He received his Bachelor's of Science in honors chemistry from the University of Alabama in 1977, and his doctorate degree in inorganic chemistry from Iowa State University in 1982. His thesis work at Iowa State focused on the synthesis, characterization, and structure of air-sensitive metal-metal bonded clusters of molybdenum and tungsten. Carlin received his training in electrochemistry as a postdoctoral fellow with Prof. Robert A. Osteryoung at the State University of New York at Buffalo.

Carlin has published more than 100 technical papers and one book chapter, and he is a co-inventor on seven U.S. patents. He has given numerous presentations including invited talks at international venues in Japan, France, Turkey, Crete, and Scotland.

Carlin was awarded the Department of the Navy Meritorious Civilian Service Medal in August 2008. In January 2001, he received Assistant Secretary of the Navy (Research, Development, & Acquisition) Awards for the Rapid Transition of Foreveready Missile Battery for Strategic System Programs and for Lithium-Ion Polymer Battery for Advanced Seal Delivery System. He was awarded the Chief of Naval Research's Award of Merit for Group Achievement in August 2000 for Superior Group Effort While Serving on the ONR Diversity Committee. Additionally, his discovery of a novel battery technology was recognized with the U.S. Air Force Materiel Command S&T Achievement Award in 1993.

THIS PAGE INTENTIONALLY LEFT BLANK

Plenary Panel 11: GAO Observations on DoD's Weapons Acquisition Portfolio, Requirements, and Acquisition Reform Efforts

Thursday, April 27, 2017	
9:30 a.m. – 11:00 a.m.	**Chair: Michael J. Sullivan,** Director, Acquisition and Sourcing Management Team, Government Accountability Office ***GAO Observations on DoD's Weapons Portfolio: Performance, Reform Implementation, and Use of Knowledge-Based Best Practices*** Desiree Cunningham, Senior Analyst, Acquisition and Sourcing Management, GAO ***Early Systems Engineering Positions Programs for Success*** Travis Masters, Assistant Director, Acquisition and Sourcing Management, GAO ***Challenges DoD Has Attracting Non-Traditional Companies to Modify Their Products for DoD's Use*** Cheryl K. Andrew, Assistant Director, Acquisition and Sourcing Management, GAO

Michael J. Sullivan—currently serves as Director, Acquisition and Sourcing Management, at the Government Accountability Office (GAO). This group has responsibility for examining the effectiveness of the DoD's acquisition and procurement practices in meeting its mission performance objectives and requirements. In addition to directing reviews of major weapon system acquisitions such as the Joint Strike Fighter, F-22, Global Hawk, and various other major weapon acquisition programs, Sullivan has developed and directs a body of work examining how the DoD can apply best practices to the nation's largest and most technically advanced weapon systems acquisition system. This work has spanned a broad range of issues critical to the successful delivery of systems, including technology development, product development, transition to production, software development, program management, requirement-setting, cost estimating, and strategic portfolio management. The findings and recommendations from this work have played a major role in the department's recent acquisition policy revisions. Most recently, Sullivan has directed the GAO's annual assessment of major weapon systems programs for Congress and the GAO's work with Congress in establishing acquisition policy reforms. His team also provides Congress with early warning on technical and management challenges facing these investments.

Sullivan has been with the GAO for 29 years. He received a bachelor's degree in political science from Indiana University and a master's degree in public administration from the School of Public and Environmental Affairs, Indiana University.

Desiree Cunningham—Senior Analyst, Acquisition and Sourcing Management, GAO

Travis Masters—Assistant Director, Acquisition and Sourcing Management, GAO

Cheryl K. Andrew—Assistant Director, Acquisition and Sourcing Management, GAO

THIS PAGE INTENTIONALLY LEFT BLANK

Panel 12. Improving Acquisition Workforce Competency

Thursday, April 27, 2017	
11:15 a.m. – 12:45 p.m.	**Chair: Reuben Pitts,** President, Lyceum Consulting, LLC ***A Description of the Defense Systems Engineering Career Competency Model*** Clifford A. Whitcomb, Naval Postgraduate School Rabia Khan, Naval Postgraduate School Dana Grambow, Office of Personnel Management José Vélez, Deputy Assistant Secretary of the Navy (RDT&E) Jessica Delgado, Deputy Assistant Secretary of the Navy (RDT&E) Corina White, Naval Postgraduate School ***Team Leader Development Needs and Competencies in the Defense Acquisition Workforce*** Stephen Trainor, Naval Postgraduate School ***Improving Workforce Professionalism: A Retrospective View of Developing Leadership Mass Through Your Staff*** Donna J. Kinnear-Seligman, Defense Acquisition University

Reuben Pitts—is the President of Lyceum Consulting. He joined the Naval Weapons Lab in Dahlgren, VA, in June 1968 after graduating from Mississippi State University with a BSME. His early career was spent in ordnance design and weapons systems. He subsequently served on the planning team to reintroduce the Navy to Wallops Island, VA, currently a multiple ship combat, over-the-water weapons testing lab for Surface Ship Combat Systems, Fighter Aircraft, and live missile firings. His outstanding service as the deployed science advisor to commander, U.S. Sixth Fleet, was recognized with the Navy's Superior Civilian Service (NSCS) Award and the Navy Science Assistance Program Science Advisor of the Year Award.

Pitts was selected to lead the technical analysis team in support of the formal JAG investigation of the downing of Iran Air Flight 655 by USS *Vincennes*, and participated in subsequent briefings to CENTCOM, the chairman of the joint chiefs, and the secretary of defense. As head, Surface Ship Program Office and Aegis program manager, Pitts was awarded a second NSCS, the James Colvard Award, and the John Adolphus Dahlgren Award (Dahlgren's highest honor) for his achievements in the fields of science, engineering, and management. Anticipating the future course of combatant surface ships, Pitts co-founded the NSWCDD Advanced Computing Technology effort, which eventually became the Aegis/DARPA-sponsored High Performance Distributed Computing Program, the world's most advanced distributed real-time computing technology effort. That effort was the foundation for the Navy's current Open Architecture Initiative. In 2003, Pitts accepted responsibility as technical director for PEO Integrated Warfare Systems (IWS), the overall technical authority for the PEO. In September of that year, he was reassigned as the major program manager for Integrated Combat Systems in the PEO. In this position, he was the program manager for the Combat Systems and Training Systems for all U.S. Navy Surface Combatants, including aircraft carriers, cruisers, destroyers, frigates, amphibious ships, and auxiliaries. In July 2006, Pitts returned to NSWCDD to form and head the Warfare Systems Department. While in this position, he maintained his personal technical involvement as the certification official for Surface Navy Combat Systems. He also served as chair of the Combat System Configuration Control Board and chair of the Mission Readiness Review for Operation Burnt Frost, the killing of inoperative satellite USA 193.

Pitts has been a guest speaker/lecturer/symposium panelist at many NAVSEA-level and DoD symposiums and conferences and at the Naval Postgraduate School, the Defense Systems Management College, and the National Defense University. For 19 years, Pitts was the sole certification authority of all Aegis Combat System computer programs for fleet use. He retired from the U.S. Civil Service in September 2008, with over 40 years of service to the Navy.

A Description of the Defense Systems Engineering Career Competency Model

Clifford A. Whitcomb—is Professor in the Systems Engineering Department at the Naval Postgraduate School in Monterey, CA. He has more than 35 years of experience in defense systems engineering. He is an INCOSE Fellow, has served on the INCOSE Board of Directors, and was a Lean Six Sigma Master Black Belt for Northrop Grumman Ship Systems. Dr. Whitcomb earned his BS in Engineering (nuclear engineering) from the University of Washington in 1984, MS degrees in Naval Engineering and Electrical Engineering and Computer Science from MIT in 1992, and his PhD in Mechanical Engineering from the University of Maryland in 1998.

Rabia H. Khan—is a Faculty Research Associate in the Systems Engineering Department at the Naval Postgraduate School, Monterey, CA. Khan's research interests include systems engineering competency (modeling and development), cognitive processing, and measuring self-efficacy within the field of systems engineering. She earned a bachelor's degree in Psychobiology from the University of California, Davis, and a master's degree in Engineering Systems from the Naval Postgraduate School.

Dana Grambow—is a Personnel Research Psychologist who joined the Leadership and Workforce Development Assessment section of the Office of Personnel Management in 2010. She has worked with numerous agencies on a variety of projects including occupational analysis, competency modeling, gap analysis, and assessment development. Dana holds a PhD in Industrial/Organizational Psychology from the University of Missouri—St. Louis.

José Vélez—is currently serving as the DASN (RDT&E) Technical Workforce Lead, shaping and defining DoD policy that directly impacts the technical workforce within the Naval Research & Development Establishment (NR&DE). He received a Bachelor of Science degree in Mechanical Engineering from the University of Puerto Rico, College of Engineering, in 1979 and a Master of Science degree in Engineering Administration from George Washington University in 1985. He is a member of the Acquisition Professionals Community and is DAWIA certified at Level III in both the Engineering and Program Management career fields.

Jessica Delgado—is in the Platforms Integration Safety Branch of the Naval Surface Warfare Center in Dahlgren, VA. She was previously the Technical Workforce Strategy Lead within the Deputy Assistant Secretary of the Navy for Research, Development, Test, and Evaluation (DASN [RDT&E]). Prior to that Delgado was the PESOH for all mortar programs for the Infantry Weapons Program Office of MARCORSYSCOM. Delgado started her career as a scientist at NSWCDD in the Concepts and Experimentation Branch in the Chemical, Biological, and Radiological division. Delgado has a Master of Science degree in Biology from the University of Puerto Rico and a Bachelor of Science degree in Microbiology from the same university.

Corina White—has expanded her professional work experience from 2007 to 2015 as a United States Navy Civilian in several unique disciplines, including research and development, aerospace engineering, and materials engineering. She has worked with the National Aeronautics and Space Administration, Naval Air Systems Command, and currently the Naval Postgraduate School. Her education includes earning a BS in Chemical Engineering from Prairie View A&M University and an MS in Systems Engineering from the Naval Postgraduate School.

Abstract

A defense-level systems engineering competency model for use in key human resource functions, such as hiring, promoting, and administering skill(s) gap assessments, and for career path modeling and development planning, has been verified for use by the DoD. The model was verified based on analysis of a survey administered by the U.S. Office of Personnel Management of 6,011 incumbents (or employees) and 1,519 supervisors in systems engineering across the Navy, Marine Corps, Army, Air Force, and the Missile Defense Agency. This paper presents a summary of the competencies and tasks of the

resulting defense systems engineering career competency model. A comparison of the competencies and tasks among the components surveyed is presented and analyzed. Conclusions are presented, along with recommendations for making the model easily and widely available for use.

Introduction

The Department of Defense (DoD) has established a competency-based approach to strategic workforce management. This approach includes assessing the critical skills and competencies needed now and in the future within the civilian workforce, along with strategies to bridge competency and skill gaps. A competency based approach supports strategic workforce planning and effective talent management. The specifications of 5 C.F.R 300A, *Employment Practices*, a federal regulations guide, require (1) a job analysis for selection and competitive promotions in federal employment, (2) compliance with the job-relatedness requirements of the *Uniform Guidelines on Employee Selection Procedures* (43FR38290), and (3) that resulting assessments target competencies required for the occupational position. The uniform guidelines are a set of principles designed to assist employers, labor organizations, employment agencies, and licensing/certification boards in complying with requirements prohibiting discriminatory employment practices.

The Deputy Assistant Secretary of the Navy (DASN) for Research, Development, Test and Evaluation (RDT&E) sponsored the development and verification of the Systems Engineering Career Competency Model (SECCM) through collaboration with the Naval Postgraduate School, Office of Personnel Management (OPM), Navy, Army, Air Force, Marine Corps, and Missile Defense Agency. The SECCM aligns with the Engineering (ENG) acquisition career field and the related ENG competency model. DASN RDT&E supported the SECCM development since the systems engineering career field does not have an occupational series (08XX), so definitions of and expectations for systems engineers vary. The newly verified SECCM provides the source for a consistent and verified definition of defense systems engineering competencies and tasks.

OPM administered the job analysis survey to 6,011 incumbents (or employees) and 1,519 supervisors across the DoD (OPM, 2016). Survey participants were asked to evaluate each competency and task on criteria such as frequency, importance, required immediately upon entry into the position, and need for training. Incumbents rated frequency of tasks and importance of competencies for themselves. Supervisors rated importance of the tasks and competencies for themselves, as well as the average importance for each task and competency for each grade level, GS-7, GS-9, GS-11, GS-12, GS-13, GS-14, and GS-15. This paper discusses the overall SECCM and analyzes the results across the Navy, Army, Air Force, and Missile Defense Agency.

For further information on the details of the development and uses of the SECCM, see Whitcomb et al. (2014), Whitcomb, White, et al. (2015, 2016), Whitcomb et al. (2016b), Khan et al. (2016), White et al. (2016), and Whitcomb et al. (2016a).

Systems Engineering Career Competency Model

Competency models in the DoD consist of competencies and tasks. Competencies are defined as measureable patterns of skills, knowledge, abilities, behaviors, and other characteristics which an individual needs to perform work. Tasks define the duties associated with an occupation, and they are linked to competencies during job analysis. The SECCM consists of 44 systems engineering competencies and 179 systems engineering tasks. This paper presents both the competencies and tasks but focuses on the competencies. Critical competencies are identified for systems engineers at the GS-07 to

GS-15 grade levels. In Table 1, the 44 SECCM competencies are categorized into four distinct units of competence, based on the current defense acquisition Engineering (ENG) Career Field Competency Model: technical management, business acumen, analytical, and professional.

Table 1. SECCM Competencies, Organized into Four ENG Model Categories

SYSTEMS ENGINEERING CAREER COMPETENCY MODEL			
Technical Management	**Business Acumen**	**Analytical**	**Professional**
Acquisition	Industry Awareness	Transition	Communication
Risk Management	Organization	Integration	Leading High Performance Teams
Requirements Management	Cost Estimating	Design Considerations	Personal Effectiveness/Peer Interaction
Configuration Management	Proposal Process	Tools and Techniques	Problem Solving
Technical Assessment	Supplier Management	Stakeholder Requirements Definition	Professional Ethics
Data Management	Negotiations	Validation	Strategic Thinking
Software Engineering Management	Cost, Pricing and Rates/Cost Management	Verification	Coaching & Mentoring
Decision Analysis	Financial Reporting and Metrics	Mission-Level Assessment	Managing Stakeholders
Interface Management	Business Strategy	Architecture Design	Mission and Results Focus
Technical Planning	Industry Motivation, Incentives, Rewards	Implementation	Sound Judgment
	Contract Negotiations	Engineering Disciplines	Continual Learning
		Requirements Analysis	

The 44 competencies and their respective descriptions are listed in Table 2.

Table 2. SECCM Competencies and Descriptions

No.	Competency	Description
1	MISSION-LEVEL ASSESSMENT	Collaborates with user community to assess mission areas end-to-end, across system and platform boundaries, to identify and close integration and interoperability (I&I) gaps in mission critical capabilities.
2	STAKEHOLDER REQUIREMENTS DEFINITION	Works with the user to establish and refine operational needs, attributes, and performance parameters based on established processes and ensures all relevant requirements and design considerations are addressed to establish a set of baseline capability requirements.
3	REQUIREMENTS ANALYSIS	Ensures the requirements derived from the stakeholder-designated capabilities are feasible and effective, and are analyzed, decomposed, and functionally detailed across the

No.	Competency	Description
		entire system.
4	ARCHITECTURE DESIGN	Creates and maintains architectural products throughout the life-cycle integrating hardware, software, and human elements; their processes; and related internal and external interfaces that meet user needs and optimize performance.
5	IMPLEMENTATION	Applies a methodical and disciplined approach for the specification, design, development, realization, technical management, operations, and/or retirement of a system.
6	INTEGRATION	Plans, manages, and executes the systems integration process to form higher-level elements and eventually the finished products.
7	VERIFICATION	Designs an evaluation strategy and process to assess the ability of the design solution to meet performance requirements and communicate capabilities, limitations, and risks.
8	VALIDATION	Designs an evaluation strategy and process to assess the ability of the design solution to meet operational capabilities (e.g., safety, suitability, effectiveness) and communicate capabilities, limitations, and risks.
9	TRANSITION	Supports operational and sustainment planning to ensure successful acceptance of the end item by the user community.
10	DESIGN CONSIDERATIONS	Assesses conformance of the design solution with policy, legal requirements, and technical tradeoffs.
11	TOOLS AND TECHNIQUES	Applies tools, techniques, and procedures to enable systems engineering practice.
12	DECISION ANALYSIS	Identifies and assesses aspects of alternative decisions (options), including the impact and implications of each, to select a course of action.
13	TECHNICAL PLANNING	Determines the scope of the technical effort required to develop, field, and sustain the system.
14	TECHNICAL ASSESSMENT	Applies formal review process, and develops and uses technical performance measures and other metrics to measure technical progress, review life-cycle costs, and assess requirements and the effectiveness of plans.
15	CONFIGURATION	Applies standard practices to establish and maintain consistency of a product or system's attributes with its

No.	Competency	Description
	MANAGEMENT	requirements, design, and operational information throughout the life-cycle, implementing configuration changes as needed.
16	REQUIREMENTS MANAGEMENT	Develops, documents, incorporates, tracks and revises requirements documents, and maintains traceability including justification for the development of and changes to requirements.
17	RISK MANAGEMENT	Provides input into and implements a Risk Management Plan encompassing risk identification, analysis, mitigation planning, mitigation plan implementation, and tracking throughout the life-cycle of the program.
18	DATA MANAGEMENT	Applies policies, procedures, and information technology to plan for, acquire, access, manage, maintain, protect, and use data of a technical nature to support the total life-cycle of the system.
19	INTERFACE MANAGEMENT	Ensures interface definition and compliance among the system elements, as well as with other systems, by implementing control processes and measures; ensures all internal and external interface requirement changes are properly documented and communicated.
20	SOFTWARE ENGINEERING MANAGEMENT	Determines software-related considerations to address architectures, requirements mapping, integration, technical data rights, cyber-security, and suitability for intended use throughout the life-cycle.
21	ACQUISITION	Applies knowledge of laws, regulations, policies, processes, and procedures related to the life-cycle management activities needed to acquire and sustain products and services.
22	PROBLEM SOLVING	Identifies problems, issues, or failures; determines accuracy and relevance of information; generates and evaluates alternative outcomes and possible solutions, and makes recommendations or decisions.
23	STRATEGIC THINKING	Formulates effective strategies, consistent with the long-term interests of the organization, to ensure the fulfillment of objectives, priorities, and plans.
24	PROFESSIONAL ETHICS	Maintains strict compliance to governing ethics and standards of conduct in engineering and business practices.

No.	Competency	Description
25	LEADING HIGH-PERFORMANCE TEAMS	Leads and builds teams by managing group processes, providing technical direction, and fostering commitment to the mission.
26	COMMUNICATION	Expresses facts and ideas both verbally and in writing taking into account the audience and nature of the information; listens to others, attends to nonverbal cues, and responds appropriately.
27	COACHING AND MENTORING	Actively participates in either providing or receiving feedback and opportunities to learn through formal and informal methods in order to develop and advance capabilities.
28	MANAGING STAKEHOLDERS	Identifies stakeholders; builds and manages effective relationships with all stakeholders; collaborates across boundaries and finds common ground with a widening range of stakeholders; utilizes contacts to build and strengthen internal support.
29	MISSION AND RESULTS FOCUS	Aligns goals and work efforts toward fulfillment of the overall organizational mission.
30	PERSONAL EFFECTIVENESS/PEER INTERACTION	Sets personal goals; displays initiative and commitment towards completing assignments; works and collaborates with peers.
31	SOUND JUDGMENT	Makes a decision or forms an opinion by identifying, discerning, and evaluating relevant information.
32	INDUSTRY AWARENESS	Applies knowledge of the defense environment and current and emerging industry capabilities to inform development and updates to acquisition strategies.
33	ORGANIZATION	Applies knowledge of how organizations are organized across products and services, one's role in the organization, and interactions among internal and external organizations to execute the systems engineering approach and strategy.
34	COST, PRICING AND RATES	Applies knowledge of cost management practices to assist in acquisition management, strategies, and technical oversight.
35	COST ESTIMATING	Applies knowledge of cost estimating requirements, methods, and key process elements to contribute to the preparation of cost estimates.

No.	Competency	Description
36	FINANCIAL REPORTING AND METRICS	Applies knowledge of financial reports and metrics to better enable program decisions.
37	BUSINESS STRATEGY	Applies knowledge of strategic planning, marketing, market research, and business development to contribute to the preparation of appropriate acquisition strategies and solicitations.
38	PROPOSAL PROCESS	Applies knowledge of the scope of work during the proposal planning and preparation process to support acquisition strategies and solicitations.
39	SUPPLIER MANAGEMENT	Applies knowledge of supply chain management to contribute to the preparation of acquisition strategies and solicitations and to provide necessary technical oversight.
40	INDUSTRY MOTIVATION, INCENTIVES, REWARDS	Applies knowledge of incentive based acquisitions, within the constraints of competition and cost, to contribute to the preparation of acquisition strategies and solicitations and to assess if technical award fee criteria are met.
41	NEGOTIATIONS	Communicates objectively with others to reach mutually acceptable solutions.
42	CONTINUAL LEARNING	Masters and applies new technical and business knowledge; recognizes own strengths and weaknesses; pursues self-development; seeks feedback from others and opportunities to master new knowledge.
43	ENGINEERING DISCIPLINES	Applies knowledge of multiple engineering disciplines and how they relate to each other and the system to achieve system performance.
44	CONTRACT NEGOTIATIONS	Applies knowledge of successful negotiations from both a government and business perspective to get the best value for the taxpayer and promote a fair profit to contribute to the preparation of appropriate acquisition strategies and solicitations.

The systems engineering tasks are listed in Table 3. More than one competency may be needed to accomplish a task. The tasks that are aligned under competencies are the ones with the most impact for the respective competency as determined by subject matter expert panels facilitated by OPM during the job analysis process. All tasks are identified as critical for systems engineering, although which tasks are critical changes with GS level.

Table 3. SECCM Tasks

No.	Task
MISSION-LEVEL ASSESSMENT	
1	Analyzes gaps between mission objectives, mission threads, existing or planned capabilities, and available funding to enable program decisions.
2	Analyzes mission-level requirements to determine if they are feasible across a program or enterprise (e.g., component, DoD, federal agencies, international coalitions).
3	Analyzes the solution space to identify potential solutions that meet mission requirements and leverage opportunities.
4	Conducts trade analysis to refine a proposed solution to meet mission requirements.
5	Contributes to the development of various scenarios for system use, functions, and performance in line with the Concept of Operations.
6	Contributes to the development of operational and top-level systems requirements that are traceable to mission-level requirements, feasible, complete, and verifiable.
7	Identifies and analyzes mission technical problems, issues, risks, and opportunities to enable informed program decisions.
STAKEHOLDER REQUIREMENTS DEFINITION	
8	Collaborates with stakeholders to set expectations and build consensus in regards to requirements throughout the system life-cycle.
9	Develops scenarios and use cases for systems that will provide services, capabilities, or platforms to end-users and other stakeholders.
10	Documents the intent, decisions, and rationale for end-user requirements to ensure traceability during the development and verification stages.
11	Analyzes capability needs, operational constraints, and technical limitations in collaboration with the stakeholders to derive system requirements and technical performance measures for system development.
12	Elicits stakeholder requirements to build a recommended list of potential system requirements.
13	Defines the constraints on a system solution that stem from existing agreements, interoperability, regulations, management decisions, and technical decisions.

No.	Task
14	Identifies the effectiveness and suitability requirements (e.g., HSI, ESOH, reliability, availability, maintainability) that correspond to anticipated operational and support scenarios and environments.
REQUIREMENTS ANALYSIS	
15	Analyzes the threat assessment to support a materiel or non-materiel solution.
16	Analyzes model, prototype, or system performance to identify and update requirements.
17	Assesses the impact of requirements changes on the solution and program.
18	Ensures end-user requirements are well-documented by collaborating with subject matter experts and other stakeholders.
19	Decomposes requirements across a system for allocation and traceability.
20	Develops specification documents for a system.
21	Prioritizes requirements for system upgrades for future enhancements in collaboration with stakeholders.
22	Establishes threshold and objective values for system requirements in collaboration with stakeholders.
23	Manages requirements for a system to include upgrades and future enhancements or pre-planned product improvements.
24	Resolves requirement conflicts in order to establish a complete, consistent, and traceable requirement set for the system of interest throughout the life-cycle of the system.
25	Defines and manages critical technical performance measures to monitor the system development.
26	Collaborates with the test and evaluation community and other stakeholders to ensure design is projected to comply with established measures of performance and measures of suitability.
27	Creates written design requirements for system performance specification.
28	Recommends revised operational requirements or design requirements to comply with government policy, regulations, and law.
ARCHITECTURE DESIGN	
29	Assesses the overall architecture to ensure it meets established requirements.

No.	Task
30	Designs architecture solution and products (e.g., DoDAF) that capture operational and systems requirements, including interfaces, interoperability, integration, and environments.
31	Creates solutions by building analytic models and conducting experiments.
32	Establishes the functional, allocated, and product baselines for use throughout the system life-cycle.
33	Formulates scalable and adaptable solutions to account for future needs.
34	Plans and executes the technical review process taking into account suitability of design attributes.
35	Manages the creation of architecture artifacts required for program integration.
36	Assesses concept feasibility to support architectural design tradeoffs.
37	Provides technical expertise and assessments during the analysis of alternatives process to determine the best materiel solution.
38	Identifies systems interfaces and interoperability concerns to achieve resolution.
IMPLEMENTATION	
39	Ensures a balanced system solution by managing the technical aspects and their impacts to or from cost and schedule throughout the life-cycle.
40	Analyzes opportunities for the reuse of existing products.
41	Provides technical input on the development strategy for acquiring, fielding, and sustaining a system.
42	Documents planning (e.g., Systems Engineering Plan, Test and Evaluation Master Plan), resource requirements, technical data, technical reviews, and analyses throughout the life-cycle.
43	Monitors manufacturing and quality assurance to identify and resolve issues, manage risks, and ensure adherence to specifications and requirements.
44	Tracks design considerations (e.g., boundaries, interfaces, standards) to ensure they are properly addressed in the technical baselines.
45	Plans for technology refresh including assessing technical readiness of proposed system changes.
46	Manages and/or oversees the manufacturing process to ensure timely and adequate implementation that is consistent and compliant with contract

No.	Task
	standards and requirements.
47	Identifies manufacturing and production process improvements to reduce life-cycle costs or increase reliability, performance, or quality.
48	Conducts or reviews manufacturing readiness assessment to baseline required industrial and manufacturing capability and maturity.
49	Verifies availability of industrial base to support critical technologies.
INTEGRATION	
50	Develops and implements an integration approach which includes identification of integration, interface, and interoperability requirements within operating conditions.
51	Plans and executes physical and functional configuration audits to verify that the as-released baseline meets requirements, and the product and its documentation align.
52	Identifies and evaluates integration and interoperability options for evolving systems, phasing out of legacy systems, or phasing in of new systems.
VERIFICATION	
53	Analyzes product verification outcomes for a system.
54	Verifies requirements traceability from the lowest to highest level of integration.
55	Develops and implements an approach/plan to verify requirements and performance using inspection, demonstration, analysis, and testing.
56	Verifies the system design meets requirements.
57	Conducts root cause analyses of problems noted during execution of the verification plan, and develops potential corrective actions to resolve.
58	Performs or oversees developmental testing for a system.
59	Prepares or reviews artifacts or evidence (e.g., test results, Plan of Actions, and Milestones) for stakeholder acceptance and certification.
VALIDATION	
60	Analyzes system validation outcomes to enable program decisions.
61	Conducts system validation activities (e.g., operational testing, limited user testing) according to the plans.

No.	Task
62	Prepares or reviews validation plans and procedures for a system, including identification of method and timing for each activity.
63	Tests system concepts and their feasibility using prototypes, builds, or experiments.
TRANSITION	
64	Analyzes the risks to successful production transition and program sustainment activities during preparation for production.
65	Develops a product transition plan for a system.
66	Conducts transition to fielding and sustainment activities according to plan.
67	Defines technical policies, processes, and procedures for an organization.
68	Coordinates with receiving sites to ensure they have the available personnel, skills, and product transition procedures to receive the end product for a system.
69	Reviews the adequacy of packaging material, handling equipment, storage facilities, and shipping services for a system.
DESIGN CONSIDERATIONS	
70	Performs safety analyses using data collection and modeling techniques.
71	Addresses reliability, maintainability, and availability as they relate to all elements of the system life-cycle.
72	Develops a safety plan for the system that complies with safety assurance strategies, policies, and standards for a system.
73	Mitigates the life-cycle cost drivers in system design to ensure a system is affordable across the life-cycle.
74	Assesses if available or imminent technology is sufficient to meet system requirements.
75	Incorporates cyber-security protection requirements during all stages of the system life-cycle.
76	Identifies and analyzes supply chain management risk areas to enable program decisions.
TOOLS AND TECHNIQUES	
77	Advises on the suitability and limitations of models and simulations.

No.	Task
78	Collects and applies real-world data for computer generated force-on-force modeling, mathematical modeling, physical modeling, scientific research, or statistical analysis.
79	Interprets the results of modeling and simulation scenarios based on current and future operational capabilities.
80	Defines the needs and the scope of modeling, simulation, and analysis activities across systems, adhering to and applying sound verification, validation, and accreditation (VV&A) practices.
81	Identifies and recommends the best value alternatives for systems.
DECISION ANALYSIS	
82	Conducts FRACAS (Failure Review and Corrective Action System) activities during the system life-cycle.
83	Provides technical input to program documentation (e.g., Work Breakdown Structure, Cost Analysis Requirements Document, life-cycle support plan) to guide program execution.
84	Provides technical expertise for the cost/benefit analysis process.
85	Defines or evaluates technical scope needed to estimate costs for a system.
86	Develops or reviews technical plans (e.g., Test and Evaluation Master Plan, Systems Engineering Plan, Information Support Plan, Configuration Management Plan) for a system to ensure integration with other organizational plans and processes.
87	Conducts trade studies to determine the most cost effective alternative.
TECHNICAL ASSESSMENT	
88	Evaluates the system and technical documentation against technical review entry and exit criteria.
89	Participates in program and milestone reviews to ensure critical technical requirements will be met.
90	Reviews contractor deliverables to ensure adherence with the contractual requirements.
91	Conducts process improvement throughout the system life-cycle.
92	Participates in independent review teams to provide unbiased technical opinion.

No.	Task
CONFIGURATION MANAGEMENT	
93	Analyzes changes to baselines to enable program decisions.
94	Conducts and maintains configuration management during the entire system life-cycle.
REQUIREMENTS MANAGEMENT	
95	Creates a tailored requirements management process based on standard systems engineering processes and key stakeholder needs to maintain a stable configuration of system and subsystem requirements.
96	Identifies, evaluates, and/or recommends changes to non-compliant technical parameters for a system.
97	Manages stakeholder requirements and maintains traceability to the sources of stakeholder need.
98	Translates and documents the system capability requirements into technical requirements for the system performance specification.
RISK MANAGEMENT	
99	Provides technical input on the Risk Management Plan encompassing risk identification, analysis, mitigation planning, mitigation plan implementation, and monitoring throughout the system life-cycle.
100	Recommends risk prioritization to support decision-making.
101	Develops or provides input on actionable risk mitigation plans or issue resolution strategies and monitoring metrics to be used across systems and programs.
DATA MANAGEMENT	
102	Provides recommendations for the identification of data rights for system technical data.
103	Develops strategies to conduct technical data management for a system.
INTERFACE MANAGEMENT	
104	Develops procedures for interface management of a system.
105	Performs interface management during all stages of the system life-cycle.
106	Identifies consequences of system changes to interfaces and interoperability.

No.	Task
107	Reviews the suitability and feasibility of interface management strategies for a system.
SOFTWARE ENGINEERING MANAGEMENT	
108	Determines software-related considerations, impacts, and risks that must be addressed as part of the systems engineering plan.
109	Evaluates the benefits and risks associated with using Commercial Off The Shelf (COTS) products.
110	Plans for or manages post-deployment operations and sustainment of software.
111	Provides input (e.g., elements of code, parameters) to software and/or system reliability models.
112	Facilitates the acquisition of software and information technology systems.
113	Verifies that the collection, migration, aggregation, and manipulation of legacy data are compatible with stakeholders' IT systems.
ACQUISITION	
114	Analyzes the impact of supportability strategies on system readiness/performance.
115	Explains technical, planning, and programmatic information in the Request for Proposal (RFP) during review, industry days, and contractor communications.
116	Specifies technical evaluation criteria for proposals to ensure acquisition program goals will be met by the selected contractor.
117	Evaluates proposals based on technical evaluation criteria to ensure acquisition program goals will be met by the selected contractor.
118	Provides technical input to the development schedule and Contract Data Requirements List (CDRL) to demonstrate progress during acquisition.
119	Contributes to the development of relationships between contractors, government, and the organization, in collaboration with integrated project teams (IPTs), to enhance the government's ability to monitor the contractor.
120	Prepares performance based work statements in accordance with procurement best practices.
PROBLEM SOLVING	
121	Defines problems at all levels (i.e., project, program, or enterprise).

No.	Task
122	Gathers information to gain greater understanding of the change (e.g., what, when, how, why) to diagnose the problem.
123	Questions assumptions and requests that are inconsistent with the mission, objectives, problems, or solutions.
STRATEGIC THINKING	
124	Anticipates new or changed demands for programs and services and seeks information to guide or take action.
125	Assesses organizational, political, operational, economic, and technical uncertainties to pinpoint opportunities that can be exploited in changing environments.
126	Collects technical, programmatic, and other historical data that is used to illustrate the evolution and change within a system.
127	Contributes to the strategic planning process by providing input on the feasibility of organizational goals.
128	Converts organization-wide strategies and policy direction into action items.
PROFESSIONAL ETHICS	
129	Complies with governing ethics and standards of conduct in engineering and business practices to ensure integrity across the acquisition life-cycle.
130	Demonstrates ethical practices by showing consistency among principles, practices, and behaviors.
131	Maintains the confidentiality of information.
132	Instills a climate of trust by demonstrating honesty and keeping commitments.
133	Integrates government ethics responsibilities with engineering and business practices.
134	Resolves acquisition-related dilemmas by prioritizing ethical values and considering how choices impact the welfare of others.
135	Takes action to stop and correct unethical behavior and practices.
LEADING HIGH-PERFORMANCE TEAMS	
136	Builds effective team performance by creating an environment of trust, respect, and commitment to mission.

No.	Task
137	Serves as an authority on technical aspects of life-cycle definitions and the implication on the project or program for other team members.
138	Assigns technical tasks or work assignments to other team members.
139	Communicates expertise, advice, and knowledge effectively for the purpose of broadening the proficiency of others and establishing cooperative relationships.
140	Creates an active network across technical groups, regulatory groups, and other stakeholders for information sharing, collaboration, and decision-making.
141	Establishes a collaborative and open work environment within the system's or program's team.
142	Leads teams by providing proactive and technical direction and motivation to ensure the proper application of systems engineering processes and the overall success of the technical management process.
143	Uses a variety of direct and indirect consensus-building techniques to overcome resistance and reach agreement on ideas, recommendations, and solutions.
144	Works with team leaders or team members to clarify team roles and responsibilities.
145	Works with team members to specify performance expectations (e.g., results, deliverables, deadlines, metrics).
COMMUNICATION	
146	Adapts communication methods and style based on the audience and the target objectives.
147	Uses a variety of media to effectively communicate information about a system.
148	Communicates complex ideas, problems, and solutions in ways that are easily understood (e.g., using examples, visualizations, analogies, mental models, animations, discovery maps, interactive displays, prototype demonstrations).
149	Facilitates an open and supportive environment through active listening, ensuring understanding, and providing and receiving constructive feedback.
150	Writes technical reports that communicate the results of a technical assessment and provide evidence-based recommendations for a system.
151	Prepares and delivers programmatic or technical briefings.

No.	Task
COACHING AND MENTORING	
152	Mentors personnel to develop their capabilities.
153	Provides training opportunities for practitioners in the field of systems engineering.
MANAGING STAKEHOLDERS	
154	Articulates shared goals, conflicting interests, and multiple views among the key stakeholder factions.
155	Builds consensus among multiple stakeholders by using a common framework of ideas and objectives.
156	Fosters relationships with key stakeholders to gain cooperation, promote openness to new ideas and recommendations, and receive feedback about priorities for systems engineering efforts.
MISSION AND RESULTS FOCUS	
157	Contributes to the creation of a shared vision and strategic goals that are aligned with the mission.
158	Prioritizes tasks based on the mission to achieve the desired results.
159	Develops and executes a systematic approach to maximize the probability of mission success.
160	Demonstrates knowledge of operational culture and mission environment through design and life-cycle planning.
PERSONAL EFFECTIVENESS/PEER INTERACTION	
161	Accepts responsibility and accountability for one's work.
162	Dedicates the appropriate time and energy to assignments or tasks to ensure no aspect of the work is neglected.
163	Encourages openness to innovative ideas from others.
164	Facilitates the resolution of conflict by employing incremental trust building strategies.
165	Modifies behavior to deal effectively with changes in the work environment.

No.	Task
166	Tailors personal interaction and facilitation approach to achieve results even when consensus is difficult to achieve due to interpersonal and organizational obstacles.
SOUND JUDGEMENT	
167	Accepts responsibility and accountability for one's decisions.
168	Makes evidence based decisions in work tasks.
169	Seeks out and uses appropriate information and subject matter expertise to make effective decisions that balance policy, systemic needs and risks, trade-offs, and creativity.
COST ESTIMATING	
170	Reviews cost estimates for subsystem elements.
171	Ensures system needs are adequately covered and properly time phased in the budget submission.
172	Evaluates resource management products to understand their implications for the system.
173	Provides technical input on the reconciliation of independent cost estimates with program office cost estimates.
174	Uses Work Breakdown Structure (WBS), Earned Value Management (EVM), or other performance tracking techniques as tools for tracking contractor performance.
SUPPLIER MANAGEMENT	
175	Provides technical evaluations of requests for modifications to contracts.
176	Ensures that system engineering best practices are considered by both contractor and government team members in the execution of the program.
CONTINUAL LEARNING	
177	Achieves and maintains certifications required for job responsibilities.
178	Maintains cognizance of evolving technology and changing engineering environments through continual learning.
179	Pursues personal development through training, certifications, and other continuous learning.

The verified SECCM competencies and tasks can be used for “high stakes” human resource (HR) functions like creating (and maintaining) position descriptions, creating job announcements, assessing job candidates, hiring, and providing a basis for employee performance assessments and ratings. For example, as a part of the hiring process, the HR specialist would work with the engineering hiring manager to create a job announcement for posting on USA Jobs. If an occupational series exists for the vacancy the HR specialist would use USA Staffing systems to access the respective 08XX series competencies and tasks that the hiring manager could use to create the job announcement. As no occupational series exists for systems engineering, the SECCM can be used as the source for required competencies and tasks for the vacancy associating them with the 0801 General Engineer series. The SECCM identifies critical competencies and tasks by GS level that facilitate defining the desired set of competencies and tasks for the GS level for the job. The SECCM has a consistent set of competencies and tasks across the DoD. OPM analyzed each component individually to gain some insight into differences in the utilization of systems engineers within the DoD.

Competency Model Analysis

The OPM SECCM analysis includes identification of the critical competencies for each component surveyed. The results for the number of critical competencies at each grade level for each component are listed in Table 4.

Table 4. SECCM Critical Competencies by Grade Level for Each Component

Grade Level	Critical Competencies			
	Navy	Army	Air Force	Missile Defense Agency
GS-07	2	1	3	4
GS-09	7	1	6	6
GS-11	6	1	19	7
GS-12	16	2	31	12
GS-13	36	39	34	23
GS-14	40	43	44	36
GS-15	43	44	44	42

This information is shown graphically in Figure 1.

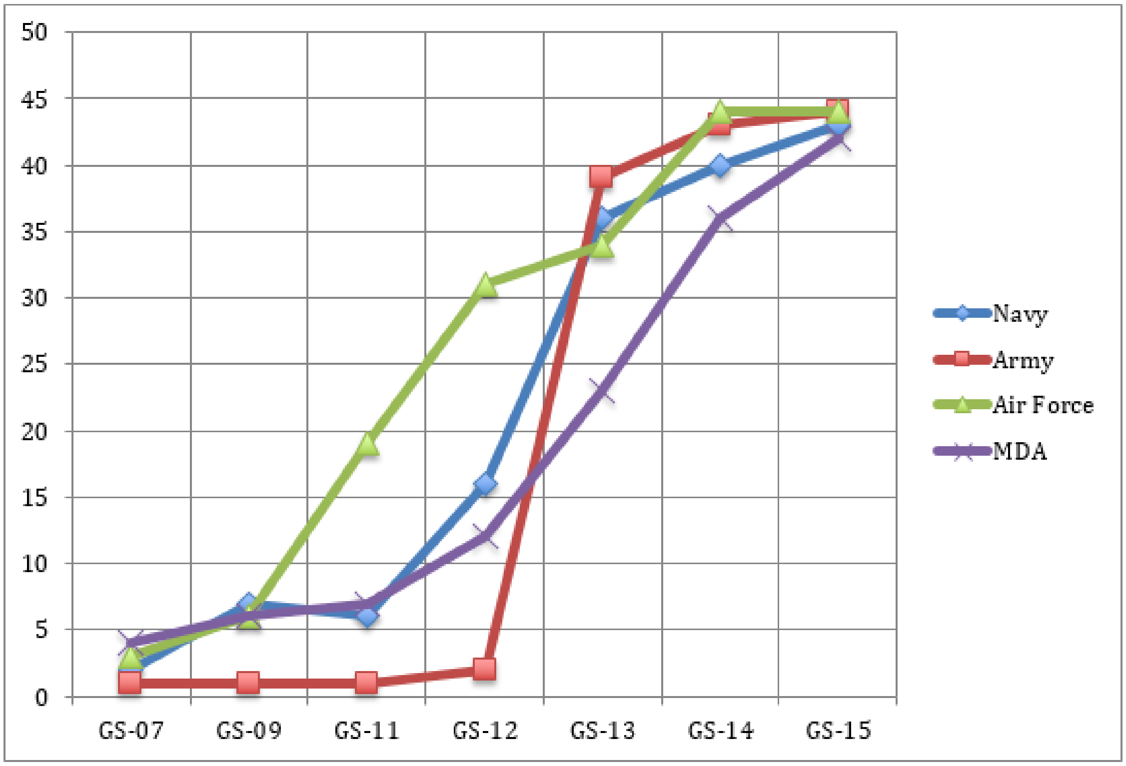


Figure 1. SECCM Critical Competencies by Grade Level for Each Component

The number of critical competencies increases with grade level for all service components. As an aside, the numbers of competencies by GS level does not indicate that all systems engineers at that level must have all of the competencies. It indicates what kinds of competencies are needed at which career levels by all systems engineers at a given level for a given component. Across all components, systems engineering competencies reach a significant number by the time an employee reaches the GS-13 level. The number of competencies required for a systems engineer reaches 34 to 39 by the GS-13 level for the Army, Navy, and Air Force and 23 for the MDA. The Army shows the most pronounced increase at GS-13, indicating that they might tend to use more senior people as their systems engineers. The Air Force shows that they expect systems engineers to develop some competencies earlier, with significant competencies by the GS-11 level. At the most senior levels at GS-14 and GS-15, almost all competencies are shown to be critical across all components.

Conclusion and Recommendations

For the first time, the DoD has a verified competency model for systems engineering. This is significant because a verified competency model is required when used for "high stakes" HR functions like creating (and maintaining) position descriptions, creating job announcements, assessing job candidates, hiring, and providing a basis for employee performance assessments and ratings. The SECCM consists of 44 systems engineering competencies and 179 systems engineering tasks. Critical competencies and tasks are identified for systems engineers at the GS-07 to GS-15 grade levels. It is recommended that the SECCM be widely distributed across the DoD using the USA Staffing system used by HR specialists. This is important until an occupational series for systems engineering can be created. In addition, it is recommended that an occupational series for systems engineering be created. The Office of the Secretary of Defense has started the process to create an occupational series for systems engineering based in part on the SECCM. This process will take time, possibly on the order of years, as the series has to be reviewed and approved for use by government agencies in addition to the DoD.

References

Khan, R., Whitcomb, C., White, C., Grambow, D., & Delgado, J. (2016). The U.S. Department of Navy's Systems Engineering Career Competency Model: Identification of proficiency levels and career path modeling. In *Proceedings of the 26th Annual INCOSE International Symposium.* Edinburgh, Scotland.

Office of Personnel Management. (2016). *Occupational analysis and workforce planning for systems engineer position* [Technical report for the Assistant Secretary of the Navy Research, Development, and Acquisition Department of the Navy].

Whitcomb, C., Delgado, J., Khan, R., Alexander, J., White, C., Grambow, D., & Walter, P. (2015). The Department of the Navy Systems Engineering Competency Model. In *Proceedings of the Twelfth Annual Acquisition Research Symposium* (Vol. 1, pp. 271–288). Retrieved from http://www.acquisitionresearch.net/files/FY2015/NPS-AM-15-004.pdf

Whitcomb, C., Khan, R. H., & White, C. (2016a). Curriculum alignment use case for competency frameworks at the Naval Postgraduate School. In *26th Annual INCOSE International Symposium*. Edinburgh, Scotland.

Whitcomb, C., White, C., Khan, R., Grambow, D., Delgado, J., & Vélez, J. (2016b, May). Update on the Department of the Navy Systems Engineering Career Competency Model. In *Proceedings of the 13th Annual Acquisition Research Symposium.* Monterey, CA: Naval Postgraduate School.

White, C., Whitcomb, C. A., Khan, R. H., Grambow, D., Delgado, J., & Vélez, J. (2016). Development of a Systems Engineering Career Competency Model for the U.S. Department of Defense. In *Proceedings of the 26th Annual INCOSE International Symposium*. Edinburgh, Scotland.

Improving Workforce Professionalism: A Retrospective View of Developing Leadership Mass Through Your Staff

Donna J. Kinnear-Seligman—is a Mission Assistance Program Analysis Manager at the Defense Acquisition University West Region. She has over 20 years of experience with developing and managing complex business knowledge applications, performing comprehensive system analyses, and conducting extensive research. Her research topics previously printed include the following:

- Challenges in Meeting OSD's Obligation & Expenditure Rate Goals (*ARJ*, Oct. 2013)
- OSD's Obligations and Expenditure Rate Goals (*AT&L*, Aug. 2013)
- Learning Organizations (*ARJ*, Apr. 2013)
- Human Capital Accelerators (*AT&L*, Mar. 2011)
- It's Time to Take the Chill Out of Cost Containment (*ARJ*, Apr. 2010)

Seligman holds an MBA (emphasis in Program Management) from the University of Management and Technology and a BS in Information Decision Systems from San Diego State University. [donna.seligman@dau.mil]

Abstract

How can the acquisition workforce train and create experience for the next generation? This research addressed how Defense Acquisition University (DAU) develops its emerging leaders among its support faculty and staff, a vital part of that equation.

Through a retrospective view of Emerging Leadership Program (ELP), the author investigated the value and impact of the program's details and its resulting effect for its graduates. To learn more, the graduates participated in a survey along with their supervisors that quantified their ELP experiences and observations, as well as confirmed any concomitant value. This research assessed the outcomes of DAU's ELP over the course of six years by focusing on various ELP activities coupled with an emphasis on behavior to reinforce the importance of improving workforce professionalism. Originally, the ELP was intended to create a development pathway. Has it?

Issue

Like any human capital development program, is the investment worth it? After completing the ELP at the Defense Acquisition University (DAU), were graduates able to influence leadership (a key performance indicator) with their new skill sets? Moreover, how many of the 58 graduates became more competitive for various leadership positions and how many actually advanced?

Introduction

Corporate universities have a long history with training programs designed to prepare their organization's future leaders—DeLoitte, the Missile Defense Agency (MDA), Disney, Motorola, InoSys, Caterpillar, MasterCard, McDonald's, to name a few. These organizations have instituted leadership programs to develop their personnel through building positive team cultures, grooming the next generation, leading through trust, and so forth. Many incorporated soft skill workshops, case studies, rotational assignments, and mentoring, etc. with many who are highly selective of their candidates. Infosys selects just 125 candidates from a list of 7,000 applicants (OriginLearning, 2014). Similarly, military organizations recognize their future leaders require formal development. Chief of Naval Operations (CNO) Admiral John M. Richardson uses the term *high-velocity learning* and recognizes the need to take advantage of the various talents and perspectives provided by the newest members of an organization (Stewart, 2016). DAU developed an Emerging

Leadership Program (ELP) that is “Developing Leadership Mass” from the bottom-up as well.

Background

How did ELP actually come about? It took root through an internal climate survey that DAU conducts every two years to learn more about what it does well and other areas that may require more attention through the use of climate survey improvement (CSI) teams. DAU established CSI teams to address any required action. In 2009, they noted a negative trend with the staff's’ job satisfaction in comparison to faculty. After a closer look, DAU staff found that staff were seeking “more recognition” and the ability to “influence decisions” in their respective workplaces. In response, the CSI team recommended that DAU develop a “Future (Emerging) Leader Program” (Seligman, 2009) and piloted an “Emerging Leadership Program” in 2011. Initial results looked promising. The graduates felt the program helped them bridge the opportunity gap that was previously missing. Today, DAU’s current ELP incarnation continues to emphasize opportunities as a cornerstone through “experience and knowledge that fosters professional and personal growth … and prepares select DAU employees for positions of increasing responsibility” (Fowler, 2015).

Emerging Leadership Program Specifics

Participation in ELP is competitive. If selected, ELP participants are exposed to a wide range of leadership competencies during a year-long program. They meet once a month virtually, and twice, face-to-face. The two face-to-face meetings are reserved for the first and last meetings. Altogether, the forums help pace the participants through various ELP program activities designed to strengthen seven core competencies:

1. Customer Service
2. Communications Skills
3. Interpersonal Skills
4. Flexibility/Adaptability
5. Problem Solving
6. Developing Others & Continuous Learning
7. Integrity & Honesty

The list below characterizes the 14 ELP components:

- Discussion Groups
- Myers-Briggs Type Indicator Assessment® (MBTI)
- Strength Deployment Inventory (SDI)
- Individual Development Plan (IDP)
- Journaling
- DDI 360 Leadership Mirror® (360)
- Team Activity
- Emotional Quotient Inventory (EQI)
- E-Learning Curriculum (e-Learning)
- Reading Report
- Soft Skill Workshops
- Mentor

- Shadow
- Final Project

ELP is designed to build greater confidence in leadership behavior through a structured process that leverages blended learning. The 14 components are used within an integrated framework of participating peers, supervisors, and mentors (formal and informal) to favorably influence graduate performance back on-the-job. For example, peers participate in 360 assessments, supervisors establish expectations, and mentors help support ELP participants in meeting their established growth goals.

Methodology

The researcher used a peer reviewed survey to collect data on eleven quantitative and qualitative questions with matrix-style format and cell groupings to expedite the time to complete (Figure 1).

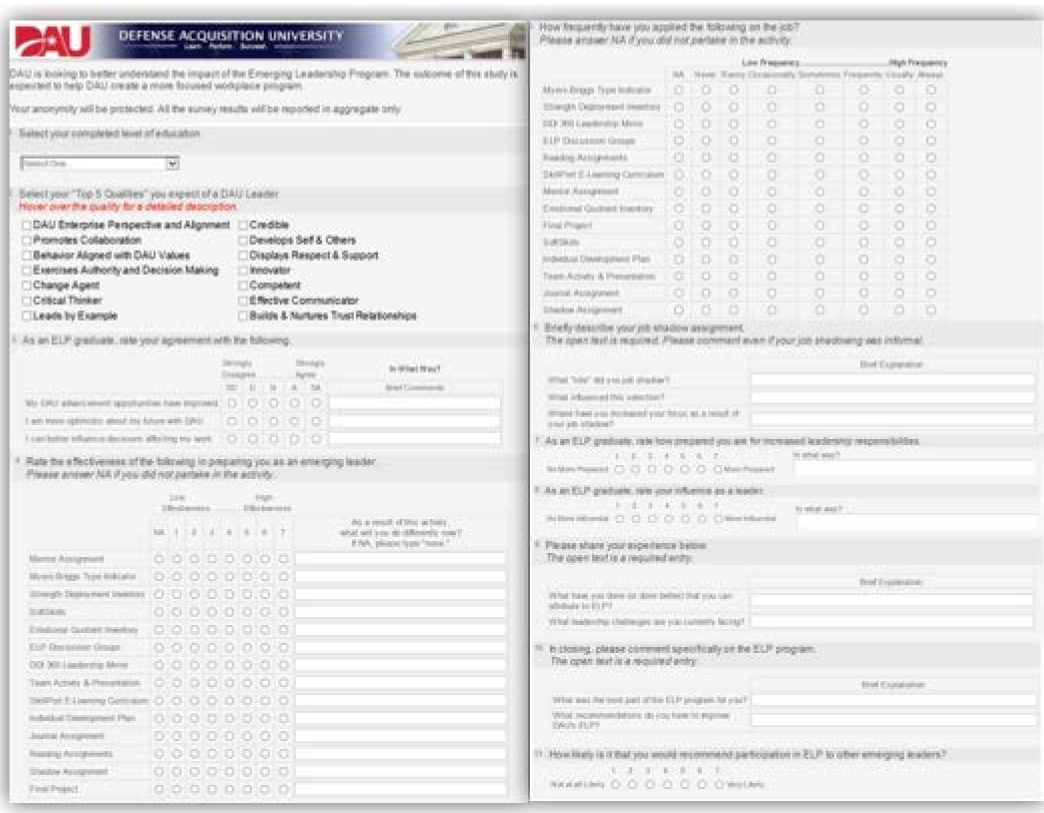


Figure 1. ELP Questionnaire

The total cohort of ELP graduates were invited to give feedback on their perceived effectiveness of the 14 ELP activities in their leadership preparation and how frequently the ELP graduates were using their newly found skills. ELP graduates were also asked to comment on how they would apply what they learned. Their supervisors were invited to participate to provide a cross-sectional perspective for this study. Because the need for a staff leadership program was triggered by a previous climate assessment, the survey also included three DAU climate survey questions for comparison to benchmarked data. The response variables were labeled using a semantic differential and then numbered with Likert-like scoring to provide a profile of the connotation.

The results were exported to Excel (Figure 2). The researcher used custom visual basic formulas to build summary arrays to display respondent groupings.

Figure 2. Summary Spreadsheet

Findings

This research confirmed the ELP's effectiveness and identified impacts through a variety of metrics based on the graduates' perceived effectiveness of the constituent ELP activities along with their supervisor's assessment of any changes in key behaviors "observed." The findings indicated a positive learning experience overall. Respondents recommended several changes to improve the effectiveness of the activities that would lead to more favorable learning outcomes including a preference for expanded leadership opportunities where they could apply their newly founded skills.

What was the single most influential factor for ELP graduates that helped them gain momentum as a future influencer/leader? It depends on the axis. For "effectiveness" it was team activities; for "frequency of use" it was soft skills closely followed by the DDI 360 Leadership Mirror®. Figure 3 highlights all 14 factors. According to the qualitative comments, the participants found this and other self-assessments including the Myers-Briggs Type Indicator Assessment (MBTI) allowed for self-awareness and self-managed change. Team activities and the unique soft skill training seemed to serve as an experiential platform to practice what they learned and helped them build stronger intellectual muscle.

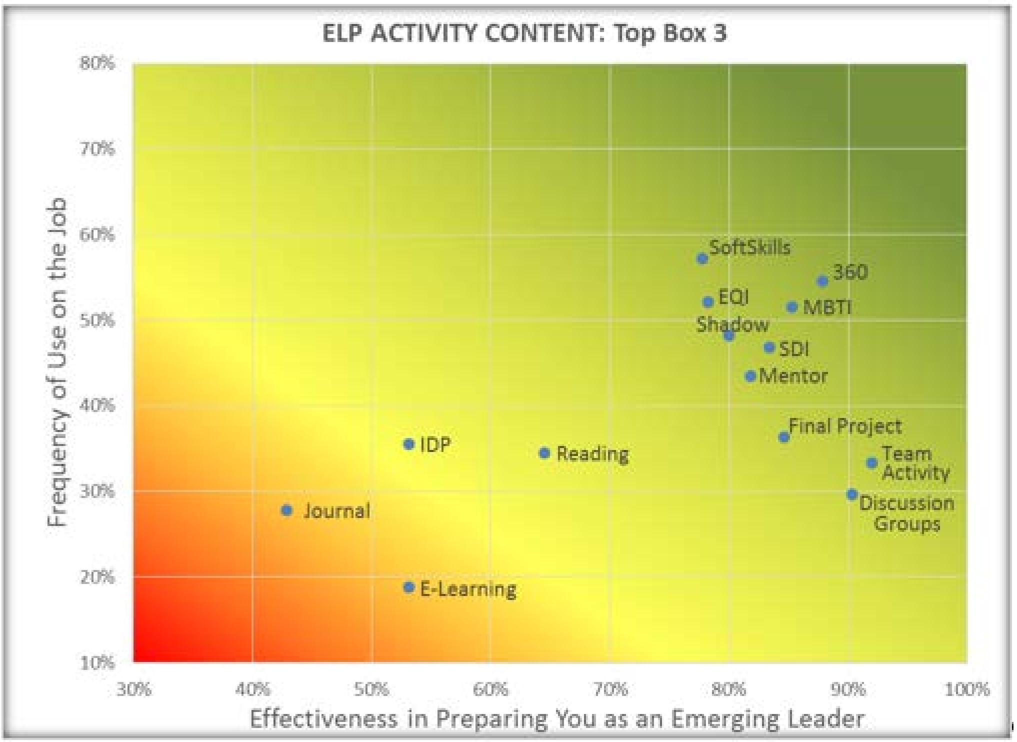


Figure 3. ELP Activity Content

Demographics

The participants in DAU's ELP were well represented in this research study, and responded at a rate of 64%. Figure 4 displays ELP year groups and their contributions to the total. With the exception of the initial ELP pilot, class sizes ranged from 9 to 13 students since the program's inception.

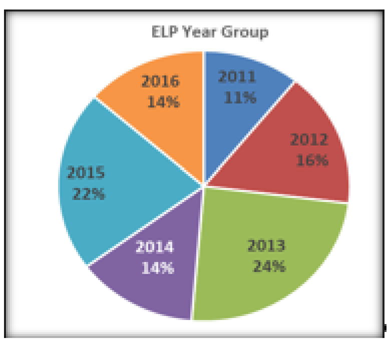


Figure 4. Contribution by ELP Year Group

Figure 5 shows how the two groups varied when comparing the respondents (n) to the total number invited (N) by year group, the gap (non-responders). Response rates also varied by generation:

- Boomers: 88%
- Gen-X: 50%
- Millennials: 71%

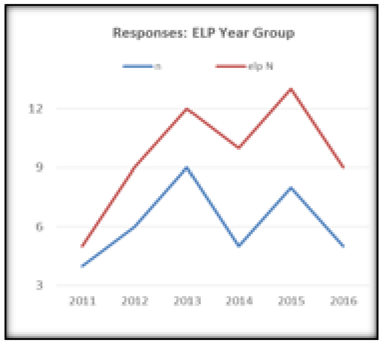


Figure 5. Year Group Response Gap

The Gen-X group was the largest non-responding group in this assessment (Figure 6). The lower response rate among Gen-X could be explained by their mistrust of technology (Erickson, 2008).

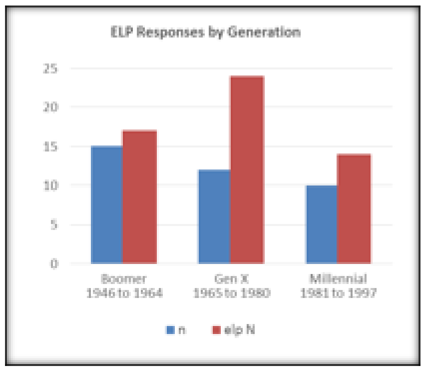


Figure 6. Response by ELP Generation

Respondents were categorized as Line Staff (admin, training techs, and specialists), Mid-Level Staff (lead specialists, management analysts, and management program analysts), or Senior Staff (designated deputy personnel or senior supervisors). The respondents also had diverse educational backgrounds. Almost a third held a master's degree while about half held bachelor's degrees. The rest either held a two-year degree or were actively seeking college credit.

Assessing the 14 ELP Activities

The 14 ELP learning activities were assessed using a top box three (TB3) methodology (i.e., totaling the responses of 5, 6, and 7 on a Likert Like scale from 1–7 and then dividing by the total respondents). Figure 7 shows the learning effectiveness of each activity. Admittedly, the effectiveness of each activity could be influenced by their frequency of use (e.g., daily, weekly, monthly, etc.). The scatterplot shows the respondents' aggregated average of the ELP attributes rated for both effectiveness and frequency.

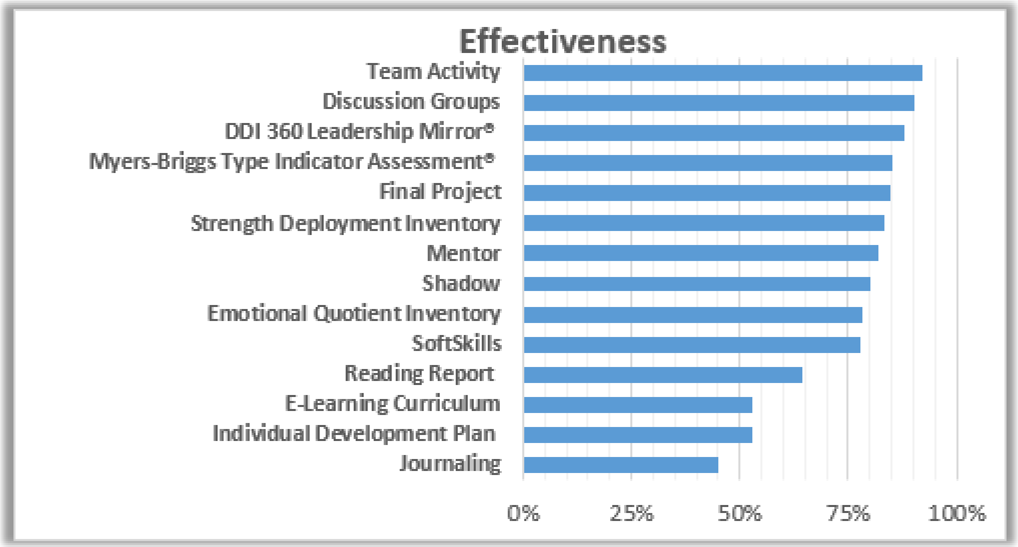


Figure 7. ELP Graduates Rate ELP Activity Effectiveness

Combining both components could suggest a tight coupling (or not) between the two. For example, the value of any particular activity with a high frequency and high effectiveness might start to wane (or not) if exercised less frequently as seen in Figure 3. To simplify the outline for the following section, the researcher kept "effectiveness" as the more influential attribute. Follow-on research could validate any changes in effectiveness for those components used more frequently (or less) over time.

Representative Comments

The following quotes represent a sampling of the respondent's comments associated with the 14 ELP activities. The comments are listed in order of the TB 3 rating (high to low) for effectiveness. Specifically, ELP graduates were asked: *"What will you do differently now?"*

Team Activity (TB: 92%) *"I will …*

… dive into projects that affect DAU as a whole."

… plan my own goals and keep others accountable to the project goals."

… stay connected with my ELP project team as a professional network."

… work better with teams to get full participation whether I'm a lead or a team member."

Discussion Groups (TB: 90%) *"I will …*

... participate more and share my ideas even when they are different."

… provide candid feedback to other members."

… get to know my counterparts from across the university."

… appreciate the power of collaboration and 'bring it' when it applies."

360 (TB: 88%) *"I will …*

… work on continuous improvement and be more self-aware in areas where I need to improve."

… improve on my effectiveness with communication skills and delegate."

… engage and work on perceptions on how I am seen by my leadership."

… continue to reflect—the 360 assessment was a profound learning moment about myself as a leader."

MBTI (TB: 85%) *"I will …*

… make more effective interactions with differing and similar personality types."

… consider adjustments needed with other personality types on how they work and interpret information."

… continue to apply MBTI and learn to better support my self-leadership."

… be more aware of other co-workers preferences so as to reduce conflict and increase group cohesion."

Final Project (TB: 85%) *"I will …*

… continue to reflect and look back at projects."

… remember this—it was rewarding to see it come together as value added."

… look for IPT participation opportunities...good vibes on my presentation."

… do this again. … I enjoyed the process of pulling together a final project."

SDI (TB: 83%) *"I will …*

… adjust my behavior and approach to conflict situations to be more effective."

… strategically approach conflict with leaders armed with knowledge of how my SDI compares to my leadership."

… be more aware of my stress reactions and make effort to better deal with daily work stressors… especially when working on teams."

… apply immediately! An informative tool and profound moment to learning about myself as a leader."

Mentor (TB: 82%) *"I will …*

… continue weekly mentor vector checks. … I have too much to work on but perspective is invaluable … learning to engage."

… find time to take on this challenge."

… reference back to my mentor when/if a situation warrants."

… utilize several suggested methods and reflect back on my mentor's insights."

Shadow (TB: 80%) *"I will …*

… request more stretch opportunities…research and/or support deep dives."

… ask for a specific project to lead now that I have had a shadow so I can exercise my new skills…I can contribute and lead from below."

… approach others with more confidence … a transformational experience."

… model my own approach from what I saw demonstrated … so much learning in the shadow!”

… look for more shadow opportunities … even if informal this is an outstanding networking opportunity and learning opportunity.”

… emulate some traits I observed as well as avoid some.”

EQI (78%) ***“I will …***

… be more patient with others.”

… handle daily situations differently.”

… show more empathy … a profound learning moment about myself as a leader.”

… treat all colleagues as humans, regardless of position or status.”

Reading (65%) ***“I will …***

… work on developing an ongoing professional reading habit.”

… look for another good leadership book.”

… use try to use and remember the skills I read…great recommended reading!”

… read additional books to continue to learn new leadership approaches and techniques.”

e-learning (53%) ***“I will …***

… look for continued e-curriculum—it was free and it was helpful.”

… use the new skills to deal with others in the circumstances I read about.”

… do more e-leaning beyond the 'requirements' to improve my self-knowledge and skill-knowledge.”

… keep soft skills on my reading list.”

IDP (53%) ***“I will …***

… better assist subordinates with setting up their IDPs now that I have done one.”

… continue IDP from a holistic perspective.”

… spend more time discussing options outside of DAU's programs that supplement the ELP learning.”

… plan my opportunities—‘need’ vs ‘want’ because I can't do everything.”

… continue to plan as I take my development seriously but in many ways I think we check the box on this.”

Journal (45%) ***“I will …***

… look for guidance on how to journal or what to record.”

… update my journal.”

Representative Comments in Review

The ELP graduates' "I will …" statements reinforced how the ELP activities may have compelled them to apply what they learned. The comments also emphasized how this program can open new behavioral pathways that could have a lasting impact. Many of the graduates spoke of seeking additional opportunities to participate in projects, IPTs, informal mentoring, and focus groups to practice and enhance their new skills. Recognizing the need to continuously learn new skills, ELP graduates seemed to recognize the importance of widening their learning apertures in the years ahead.

An Assessment of Impact

What matters is generally measured. With ELP, what is a suitable measure of effectiveness for the program as a whole? Benchmarking against DAU's internal climate survey (collected on a Likert Like scale of 5) appeared to be a sound approach. Hence, the researcher used DAU Climate Survey questions to measure impact of ELP after graduation from the year-long program. ELP graduates offered candid comments—some very provoking—as described below.

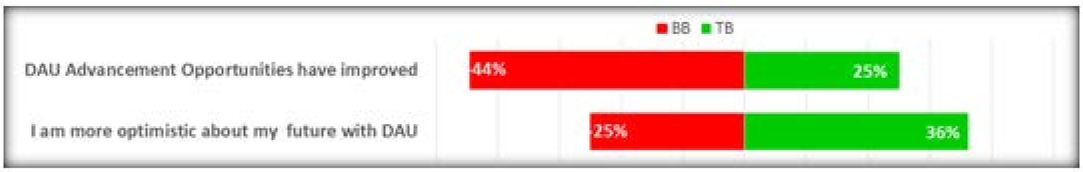


Figure 8. ELP Graduate Responses to Climate Survey Questions

DAU Advancement Opportunities and ELP Optimism (Figure 8)

Bottom Box: Have DAU Advancement Opportunities improved? With 44% in the bottom box, the majority felt opportunities had not improved; 25% of the respondents indicated less optimism with their futures at DAU. ELP, as the name suggests, is a program to develop emerging leaders. But what do ELP graduates actually emerge to? Several respondents pointed out that staff advancement is limited, "dismal at best," and "rare." Initial discussions at the program outset made it clear that there are few advancement opportunities at DAU for staff. A few respondents indicated they were encouraged to "move on" if they were seeking a leadership position—and they did. Over 10% of the graduates have left DAU for advancement positions.

Top Box: Why did some ELP graduates confirm improvements in their advancement opportunities when the data did not appear to support this ranking? While many respondents expressed disappointment that leadership and supervisory positions appeared to be fading, others stated a more optimistic view. They said they were "waiting it out" or "hanging in there" for leaders to retire or "move on." While the optimists (where one respondent said "attitude is everything") may not see a clear path for advancement (i.e., "it isn't evident"), they stressed the power of influence as a surrogate leadership strength. Several respondents noted that cross departmental and collaborative projects were especially useful. It gave them more opportunities to influence outcomes.

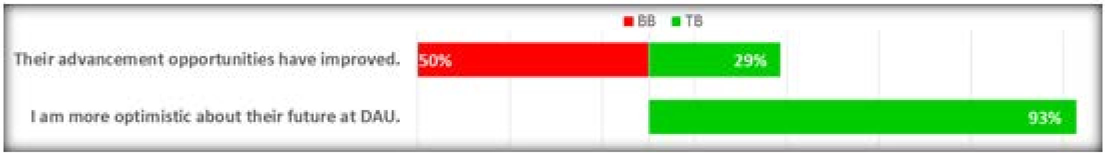


Figure 9. DAU Climate Survey Questions Answered by ELP Graduate's Supervisors

Supervisor's Speak

Supervisors of ELP graduates were asked the same questions from the climate survey (Figure 9). While both ELP graduates and their supervisors responded similarly to DAU advancement opportunities for emerging leaders, they had distinctly different levels of optimism about an ELP graduate's future. Supervisors viewed ELP graduates' future impact at DAU with 93% in TB3—significantly higher than ELP graduate responses (i.e., 57%). When asked to comment "in what way," supervisors indicated their ELP graduates seemed to possess broader perspectives of DAU. One supervisor noticed his/her ELP graduates came back to the job with a better understanding how to articulate their end-of-year contributions and how those contributions supported DAU Strategic Goals. Supervisor comments below (with ELP graduate names protected) reinforced other visible gains:

> *"xxxxx more actively seeks opportunities"*
>
> *"The experience has opened xxxx's view outside beyond her immediate environment"*
>
> *"xxxx has a greater sense of inclusion and is taking ownership."*
>
> *"The program is a plus for their evaluations."*
>
> *"I have seen a notable change in most who have graduated."*
>
> *"I see a LOT more confidence in xxxxxxx these days."*
>
> *"I see growth in leadership skills and think ELP also provided a great internal networking opportunity that will facilitate future contributions."*
>
> *"xxxx having open communication with others in the program is of value to obtain goals.*

Figure 10. DAU Climate Survey Questions Answered by Both ELP Graduates and Their Supervisors

Influencing Decisions at Work (Figure 10)

Bottom Box: Very few respondents commented in the bottom box. One respondent noted "after getting to know the other ELP participants, it is clear that influence largely depends on the daily duties of the staff member." Another said, "I'm just a worker bee …" Perhaps, seeing influence as an inherent feature of a particular leadership position is holding some ELP graduates back who could otherwise be influential. One ELP graduate conveyed a more pessimistic view and said "almost everyone here has been through the program" and none were "selected for ELP" because they were potential leaders."

Top Box: The ELP graduates scoring in the TB box (64%) appeared to be more comfortable with everything from communication to making decisions. They noted their communication skills had improved and they even understood themselves better. They also felt they developed a stronger voice along with an increased ability (and responsibility) to use it for the team or project to succeed. One respondent said "accept it—there is potential to lead from below." Others said "my boss has confidence in me to perform my duties above my duty description" and "I could influence decisions about my work … what I learned is helping me to be a better asset to DAU." Even others said they learned new skills to help them act more assertively and make a bigger difference.

Supervisor's Speak

Can ELP graduates better influence key decisions? With 100% in TB3, supervisors indicated their ELP graduates could better influence key decisions. Supervisor comments listed below confirmed several significant improvements:

> *"xxx has improved confidence and better communication skills, this combined with a better understanding of DAU culture all leads to better decision making."*
>
> *"I think most of the graduates are still a work in progress; but I think the more we can keep them involved the more they will excel and influence not just their work but work of the University."*
>
> *"xxxx has better insight into the organization and more useful contributions."*
>
> *"xxx better understands and is applying the power of influence having ascended to a more senior leadership position."*
>
> *"The program does a great job at promoting a team environment and forge lasting professional relationships. My staff reaches out to a network of thinking partners and comes to me with more actionable and relevant solutions."*
>
> *"My ELP graduate has more confidence in themselves which encourages the sharing of ideas and ultimately influence."*

Influence as a Leading Indicator for Results

ELP graduates said "with the right skill and attitude, you can influence and win confidence to make an impact." Research suggests that a leader without influence is not a very effective leader. Others say, "An Influencer is a Leader. … They challenge processes." In that context, ELP graduates who learn to influence despite certain hurdles can indeed lead up and across, and ultimately achieve some the same outcomes as leading down. The formula depicting influence in Figure 11 was adapted from a leadership blog (Rockwell, 2016) and captures the fundamentals of influence.

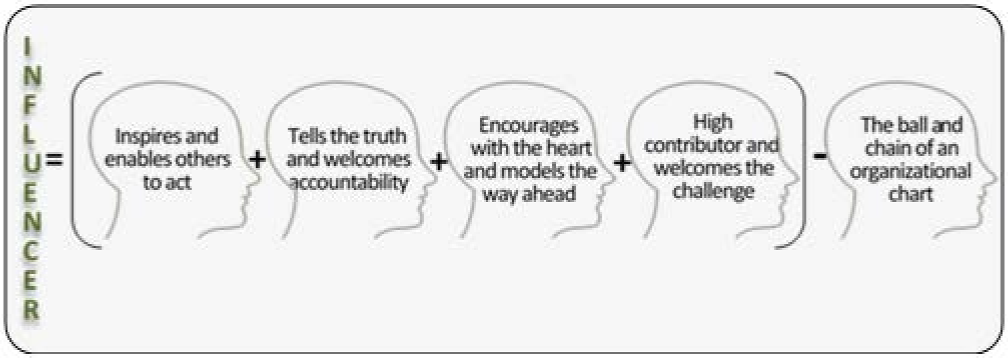


Figure 11. Influencer Model

After looking closely at all the data, did ELP create any results for the DAU, organizationally? What leading indicators or predictors suggest that behavioral changes are on track to produce desired results (Kirkpatrick & Kayser Kirkpatrick, 2016)? ELP supervisors were asked what their ELP graduates had done on the job that they could attribute to ELP. Any observed behavioral changes could serve as a bellwether of things to come.

> *"She has improved confidence in taking on challenges."*
>
> *"Communication and negotiation skills have greatly increased.*
>
> *"My ELP graduates are taking on more leadership responsibilities."*
>
> *"xxxx filled a leadership void created when a PM left and he did such a great job that the position was not back filled."*
>
> *"My ELP graduates have learned to influence my behavior to the betterment of our business unit."*
>
> *"My ELP graduate has LED several key projects in the West Region; thinks more strategically; spends more time developing herself; asks more questions of leadership."*
>
> *"Two of my staff who have graduated from ELP have been promoted and one has taken a very active role in mentoring our pathways students."*
>
> *"I noticed that xxxxx now not only actively seeks leadership opportunities but also does a great job leading teams."*
>
> *"The program broadened their perspective and their ability to see the big picture and connect the dots … a change in mindset and attitude which will go a long way."*
>
> *"My ELP graduates are taking on more lead roles for projects and have become invested in their own success."*

Leadership Qualities: What Matters and the Gaps

Supervisory data confirmed ELP Graduate contributions and their ability to influence outcomes. However, because influence can be complex to measure, ELP graduates and their supervisors were asked, "What made leaders effective?" Specifically, ELP graduates were asked to identify the leadership qualities they expect of a leader while supervisors were asked to identify the five leadership qualities they expect of ELP graduates. Figure 12 compares “what mattered most” for both groups along with their fluctuations. The companion comments provided several “perception gaps.”

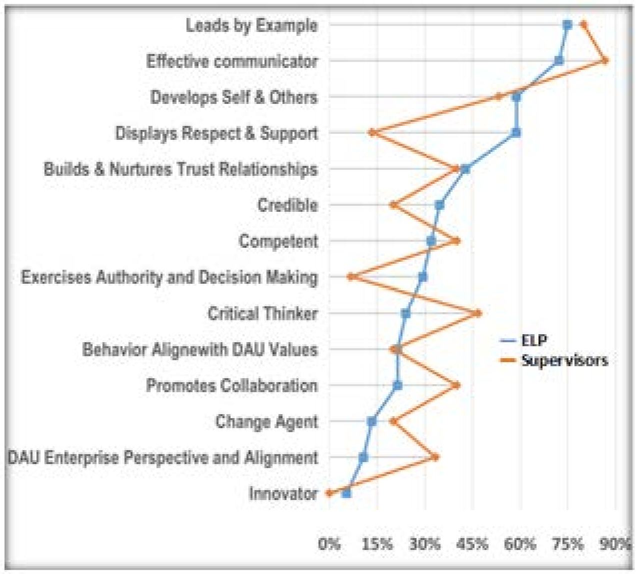


Figure 12. ELP Graduates and Supervisors Asked: What Makes Leaders Effective?

What Matters Most

While no single quality can be attributed to successful leadership, the awareness of all these characteristics can give a strong grounding for ELP graduates to influence outcomes. The characteristic that promote more influence could be the one that supervisors selected the most—communication. ELP graduates also placed high value in communication.

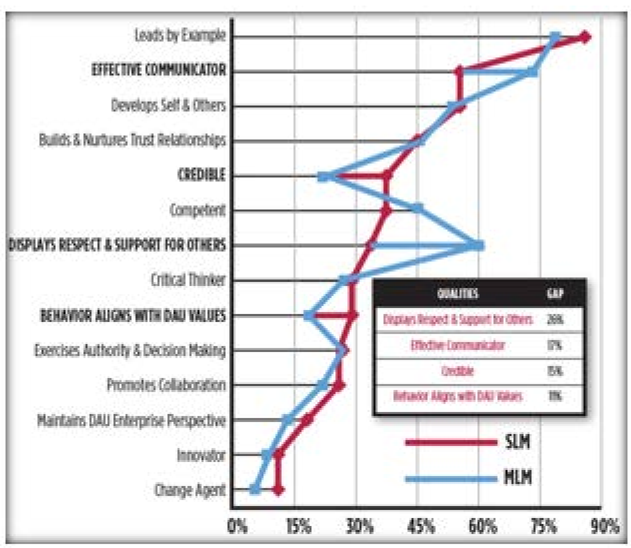


Figure 13. SLMS and MLMS Asked: What Matters Most in a Leader?

Perception Gaps

The largest gap between the groups (i.e., almost 50%) centered on Displaying Respect & Support. Only 13% of the supervisors selected this quality as the top five for the graduates while almost 60% of the ELP graduates felt it served as a much higher quality expected of a leader. Based on this aggregate level the gap, ELP graduates would be well served if they better understood their supervisor's expectations, especially in both critical thinking and collaboration.

The High Flying Leadership Qualities

Interestingly enough, the ELP graduate responses were consistent with the results of a separate study on DAU's MLM (mid-level leadership) responses (Tremaine, 2016). Tremaine noted that demographic factors easily influenced the leadership qualities. Both studies found the largest gap to be "Displaying Respect" where subordinate groups placed significantly more emphasis on it.

So what does this all mean? Are "influence," "displaying respect," and "outcomes" linked in some way? What is the cost of not displaying respect and what are the gains? A Harvard Business Review (HBR) study found that "no other leader behavior had a bigger effect on employees across the outcomes measured" (Porath, 2014)." In their study of nearly 20,000 employees, HBR noted about half didn't feel respect was displayed by their leaders. Moreover, negativity was "contagious." The other half who said they were treated with respect reported high job satisfaction, high job significance, high organizational trust, more engagement, and more likely to stay with their organization. Doug Conant (former Campbell's Soup CEO) turned Campbell's Soup around and set performance records (5 fold on the S&P ratings) largely by showing employee's respect. Stephen Covey (n.d.) reinforces respect by placing it second on his Leadership List.

What ELP Graduates Said About Learning in Their Ongoing Job Shadows

The ELP graduates shadowed DAU Regional Deans, Associate Deans, Operational Directors, Information Managers, Industry Chair, and various Directors, Department Chairs, and so forth. In some cases, they even shadowed leaders outside DAU including NAVAIR, the Air Force Institute of Technology, and other senior Learning Officers. Their experiences were captured in the following representative sampling.

> *My job shadow gave me a better appreciation for relationships and the need to stay in touch with customers to better meet their needs and keep collaborative attitude and communication lines open. I have better insight on customer relationships … this is a new perspective for me on what DAU does with our ALM that would not have happened without my ELP experience.*
>
> *I was able to observe and participate in discussions at our most senior levels. I have more insight and had a unique hands on experience.*
>
> *I learned to not get overly attached to my own ideas when they are challenged. … I've been practicing and have improved at keeping my mind open to other's input.*
>
> *Although this leader was from another organization, I saw how differently he applied himself with assertiveness and in communication. I will be increasing my focus on professional development to improve my ability to apply soft-skills.*
>
> *I wanted to experience how to handle tough situations, especially initial outreach engagements. I shadowed several seasoned experts and watched*

how they each engages with customers. My self-confidence quadrupled. I said to myself, "I can do that" and decided to be the staff example and move into unchartered territory (e.g., projects not typically assigned to staff) to show we are a key resources … we can lead and influence outcomes.

I saw what the next level might be like when it came to the daily grind. I am now more prepared and maybe less idealistic on thinking management is an easy job. I was able to witness just how much of the job is simply about solid communication with employees and customers via email, meetings, phone calls, and support of internal business processes. I am more purposeful now in my own communications.

Reflected on the positive and less effective qualities of each supervisor. For me personally, I am more reflective before making decisions and I am thinking more critically about all my work interactions so they stay positive and productive.

I have increased my focus on improving my decision making skills.

I see the value in mentoring others and having a mentor. I plan to assist others more and be more willing to ask for help.

I saw how she narrows her focus on doing less and targets the high impact projects but delegates tasks. I plan to find more guidance to help build this skill by encorporating a coach or a mentor to help me stay focused on high impact. … this is hard because I am staff I am expected to take care of work at the task level.

She is a well-respected leader and a great example of a humble leader. My own leadership style has been greatly influenced by spending time with her.

I like her individual personality. … I learned and plan to be more compassionate.

I liked the way she treated her employees … really listened to them

I learned a great deal in observing the leadership styles of the previous Dean and the new Dean.

Conclusion

After collecting and analyzing the data, DAU is in a better position to substantiate the required adjustments to the ELO program and keep it relevant, challenging, and serve as platform to further develop staff. The components that the respondents found very useful don't need much tweaking; the ones that showed less value including course readings, IDP, Journal, and e-learning could all use a boost. Nonetheless, 81% of the ELP graduates felt ELP prepared them for increased leadership responsibilities and gave them more useful tools to influence decisions. More importantly, they became more conscious of the leadership dynamic. As one respondent said, "Leaders make mistakes too, but it's how they communicate and take accountability that makes them stronger leaders and real influencers." Another respondent echoed a similar view and said, if you aspire to be a leader, take a hard look at how you influence and begin to develop those skills. Shadow assignments, self-assessments, and soft skills were all rated as skillsets used more frequently, but what needs to happen to make them all more effective? Some of the respondents felt cohorts needed to continue after graduation to continue to focus on their development. Supervisors saw marked gains in their ELP graduates' confidence and ability to influence. The year-long time investment seemed well-worth it given the observed results. The actual behavioral changes they noted in the workplace speaks strongly to the ELP's efficacy. As an enterprise, DAU heavily depends on its professional staff to meet its mission. Without them, DAU would be hard-pressed to provide a global learning environment to develop qualified acquisition, requirements, and contingency professionals who deliver and sustain effective and affordable warfighting capabilities. As a janitor so aptly replied when President John F. Kennedy asked the janitor what he did at NASA, the janitor replied, "I'm helping put a man on the moon!" (Nemo, 2014). DAU's investment in the development of their staff through ELP can do the something similar for the DAW. For ELP, training was just one component. On-the-job behaviors were more telling and they got a noticeable boost through an integrated framework of participating peers, supervisors, and mentors (formal and informal). In practice, this framework also served to monitor, reinforce, encourage, and reward ELP participants. It also helped develop a shared commitment to sustain greater confidence for the ELP graduates where they learned by doing—via "blended performance learning" (Kirkpatrick, 2016). The integrated framework helped ELP graduates stay ahead of the learning curve. In a recent poll, 41% of the ELP graduates said they were promoted and 10% left DAU for leadership positions. Although they were not promoted, an additional 13% said they have been assigned increased leadership responsibilities. It's safe to say that other organizations who are intent on developing and improving the professionalism of their workforce, including staff, can achieve measurable gains by instituting something similar to DAU's ELP. As a DAU ELP graduate, it did for me.

Recommendations

- Determine the critical behaviors that will be required for your organization's emerging leaders and build a program around it. Invest early in a formal and competitive leadership development program that has clear objectives, expectations, and is communicated enterprise-wide.
- During the leadership development program:
    - Institute periodic self-assessments benchmarked against peer groups to confirm growth areas and uncover blind spots that may require more developmental attention.
    - Leverage the value and safety of cohort teams for their collaborative contributions to innovative learning.
    - Assign team leaders to group projects so they can exercise a wide range of leadership skills through experiential learning.
    - Recognize the significance of developing "influencers" even though individuals in the development program may not be occupying formal leadership positions.
    - Seek feedback from both the participants and their supervisors throughout the participants' development journey, and share the results in a timely manner.
    - Adjust program content as required to keep it relevant and robust.

References

Covey, F. L. (n.d.). *Speed of trust: The 13 behaviors of high-trust leaders*. Retrieved from http://www.franklincovey.com/facilitator/minisessions/handouts/13_Behaviors_MiniSession_Handout.pdf

Erickson, T. (2008). 10 reasons Gen Xers are unhappy at work. *Harvard Business Review*, 1. Retrieved from https://hbr.org/2008/05/10-reasons-gen-xers-are-unhappy

Fowler, D. (2015, October 19). Emerging leader program. *The Emerging Leader Program Guide*. Fort Belvoir, VA: Defense Acquisition University. Retrieved from https://daunet.dau.mil/Personnel/SitePages/Emerging%20Leader%20Program.aspx

Kirkpatrick, J. (2016, September 7). *Kirkpatrick silver and bronze training*. Lecture presented at Kirkpatrick Four Levels® Evaluation Certification Program—Bronze and Silver Level, Fort Belvoir, VA.

Kirkpatrick, J. D., & Kayser Kirkpatrick, W. (2016). *Four levels of training evaluation.* Alexandria, VA: ATD.

Nemo, J. (2014). What a NASA janitor can teach us about living a bigger life. *How to: Growth strategies.* Retrieved from http://www.bizjournals.com/bizjournals/how-to/growth-strategies/2014/12/what-a-nasa-janitor-can-teach-us.html

OriginLearning (2014, September 8). Infosys University: The largest corporate technical university in Asia. Retrieved from https://originlearning.in/Top5Learning/beta/index.php/2014-09-11-13-51-07/corporate-universities/471-infosys-university-the-largest-corporate-technical-university-in-asia

Porath, C. (2014, November 19). Motivating people: Half of employees don't feel respected by their bosses. Retrieved from https://hbr.org/2014/11/half-of-employees-dont-feel-respected-by-their-bosses

Rockwell, D. (2015, December 29). *The exponential leadership goal for 2016* [Blog]. Retrieved from https://leadershipfreak.wordpress.com/2015/12/29/the-exponential-leadership-goal-for-2016/

Seligman, D. (2009). *Climate survey IPT: Staff*. VA: DAU.

Stewart, K. (2016, June 30). Chief of naval operations shares his vision for high velocity learning. *Naval Post Graduate School: University News Archive.* Monterey, CA: Naval Post Graduate School. Retrieved from http://www.nps.edu/About/News/Chief-of-Naval-Operations-Shares-His-Vision-for-High-Velocity-Learning.html

Tremaine, R. L. (2016, April). The high flying leadership qualities: What matters the most? *Defense Acquisition University Acquisition Review Journal, 77*(2), 121–14

Panel 13. Containing Program Lifecycle Costs

Thursday, April 27, 2017	
11:15 a.m. – 12:45 p.m.	**Chair: Gary R. Bliss,** Director, Performance Assessments and Root Cause Analyses (PARCA), Office of the Assistant Secretary of Defense (Acquisition) **Discussant: John Dillard, COL, USA (Ret.),** Naval Postgraduate School ***Application of Should Cost Methodologies to Address Savings Across the Acquisition Life Cycle*** Steve Breitenstein, Air Force Institute of Technology Bruce Johnson, Air Force Institute of Technology Jason Borchers, Air Force Institute of Technology Mark Caudle, Air Force Institute of Technology Lt Col E. Jay Kilpatrick, Air Force Institute of Technology ***Creating a "Should Cost" Culture Through Opportunity Management*** David Riel, Defense Acquisition University

Gary R. Bliss—is Director, Performance Assessments and Root Cause Analyses (PARCA), in the Office of the Assistant Secretary of Defense for Acquisition. PARCA carries out performance assessments of Major Defense Acquisitions Programs (MDAPs) and conducts root cause analyses for MDAPs with Nunn-McCurdy breach status or when requested by senior Department of Defense (DoD) officials.

Bliss previously held the position of Deputy Director, Enterprise Information and Office of the Secretary of Defense (OSD) Studies in the Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics (AT&L). His responsibilities included oversight of the five OSD-funded Federally Funded Research and Development Centers and the OSD's university research program, as well as review and development of innovations to overhaul the AT&L enterprise management systems.

Earlier in his career, Bliss served 13 years as the Director of Office of the Director, Program Analysis & Evaluation Weapon System Cost Analysis Division (WSCAD). WSCAD's 10 staff members constitute one of the two offices dedicated to OSD Cost Analysis Improvement Group (CAIG) functions, and is responsible for the preparation of independent development and procurement cost estimates for major systems that range from munitions (e.g., tactical missiles) through platforms (e.g., helicopters, submarines, fighter aircraft, tanks, etc.). As such, Bliss has been a key player in the DoD's most important system decisions by the Services, OSD, and Congress.

Generally recognized in both industry and government as a leading authority on the economics of defense procurement, Bliss has an established track record in institutional reform and enterprise reengineering. He is an experienced lecturer, often speaking to varied audiences on such topics as

- Management information system governance and reengineering
- Manufacturing enterprise reengineering
- Acquisition institutional reform

Owing to this expertise, Bliss has been hosted by the governments of Australia, Taiwan, Japan, and the United Kingdom to lecture their staffs on matters of defense acquisition.

Bliss has a BA, Mathematics and Economics (Highest Honors in Economics), from College of William and Mary.

John T. Dillard, COL, USA (Ret.)—is the Academic Associate for Systems Acquisition Management at the Graduate School of Business and Public Policy, Naval Postgraduate School (NPS) in Monterey, CA. He began his Army service as a Ranger-qualified infantryman and master parachutist, serving in the 1st Infantry and 82nd Airborne divisions, and joined the NPS faculty in 2001 upon retiring from the Army after 26 years of service. He spent 16 of those years in acquisition, most recently as commander of the Defense Contract Management Agency, Long Island, NY. He has also served on the faculty of the U.S. Army War College and as an adjunct professor of project management for the University of California, Santa Cruz. He holds an MS in systems management from the University of Southern California and is a distinguished military graduate of the University of Tennessee at Chattanooga with a BA in biological sciences.

Application of Should Cost Methodologies to Address Savings Across the Acquisition Life Cycle

Steve Breitenstein—is an Instructor at the Air Force Institute of Technology (AFIT) teaching Should Cost Management for the School of Systems and Logistics. He holds a BBA from Ohio University and an MBA from Kent State University. Breitenstein spent 23 years in industry as a program manager and business analyst and is a certified Project Management Professional (PMP). [steve.breitenstein@afit.edu]

Bruce Johnson—is an Instructor at the Air Force Institute of Technology (AFIT) where he teaches a variety of cost estimating and financial management topics for the School of Systems and Logistics. He holds a BS from Oregon State University and an MA from Webster University. Johnson spent 21 years as an Air Force Financial Manager and Cost Estimator and seven years as an instructor at AFIT. He is a Certified Cost Estimator/Analyst. [bruce.johnson@afit.edu]

Jason Borchers—is a Course Director and Instructor at the Air Force Institute of Technology (AFIT) where he teaches a variety of cost estimating and financial management topics for the School of Systems and Logistics. He holds a BS in Accounting from Wright State University and an MS in Acquisition Management from AFIT. Borchers has spent his entire career in the Budgeting and Cost Analysis career fields to include over 20 years as an active duty Air Force officer and more than four years as a Course Director/Instructor at AFIT. He is a Certified Defense Financial Manager - Acquisition (CDFM-A). [jason.borchers@afit.edu]

Mark Caudle—is the Project Management Team Lead and an Instructor at the Air Force Institute of Technology (AFIT) who teaches and consults in project management courses and workshops for the Air Force acquisition workforce. He holds a PhD from Ohio State University, an MS from the Air Force Institute of Technology, and a BA from the University of North Carolina at Chapel Hill. Dr. Caudle spent almost 22 years as an Air Force Acquisition Program Manager, and he has taught at AFIT as a civil servant for the past 10 years. [mark.caudle@afit.edu]

Lt Col Edwin "Jay" Kilpatrick, USAF—is currently a Course Director at the Air Force Institute of Technology (AFIT). He provides defense-focused professional continuing education, service, and research for the Air Force acquisition and logistics communities. He has served in a variety of training and acquisition assignments in support of the T-38, F-22, F-35, F-16, KC-10, C-9, and V-22 aircraft. Lieutenant Colonel Kilpatrick was commissioned in 1996 upon graduation from The Citadel, Military College of South Carolina, with a BS in Business Administration, and also has attained an MA from Air University in Military Operational Art and Science. [edwin.kilpatrick@afit.edu]

Abstract

The goal of this research is to examine the Air Force Should Cost Methodology and how effective it is at influencing programs to document and implement Should Cost Initiatives (SCIs) and savings opportunities. The Air Force produced a comprehensive program that includes a workshop where AFIT facilitates a seven-step process to discover SCIs for attending programs. The goal of the Workshop is for each program to develop at least one SCI. AFIT is not the expert on the individual programs; the attendees are, and AFIT guides the Workshop instead of directing specific methods. As a result, since the inception of the AFIT-led Should Cost Workshop in March 2016, 31 IPTs have attended with 89 SCIs identified and $1.039 billion reported by the IPTs as potential savings over the upcoming 1–3 year timeframe. Many lessons learned have been documented, resulting in several recommendations to improve the Workshop.

Full Abstract

The goal of this research is to examine the Air Force Should Cost Methodology and how effective it is at influencing programs to document and implement Should Cost Initiatives (SCIs) and savings opportunities. This methodology is in accordance with the

Department of Defense (DoD) Better Buying Power (BBP) initiative that requires programs to actively manage costs through the careful assessment of the contributing drivers of cost across a program, identification of goals for cost reduction, and implementation of specific efforts designed to achieve those cost reductions. The Air Force decided to produce a program in two phases. Phase 1 was to develop an asynchronous distance-learning course for the Air Force acquisition workforce to educate them on and provide wide exposure to the current Air Force policy on Should Cost. This resulted in the development of the distance learning course SYS 190 Air Force Should Cost Fundamentals. With this foundational knowledge now available to the entire Air Force acquisition workforce, AFIT moved to Phase 2, which was to develop a facilitated the Workshop where AFIT faculty SMEs would help guide complete integrated project teams (IPTs) to follow the seven-step process to discover SCIs on their projects.

The Air Force Should Cost Workshop (WKSP 0656) is sponsored by the Secretary of the Air Force/Acquisitions (SAF/AQX) and is designed to be delivered to IPTs rather than individual students. Each team consists of representatives from program management, contracting, financial management, engineering, cost analysis, logistics, and other subject matter experts (SMEs) that form the core IPT. The goal of the Workshop is for each IPT to develop at least one SCI for their program. A key strength of the Workshop is AFIT personnel act as facilitators rather than instructors. AFIT is not the expert on the individual IPTs; the attendees are, and AFIT guides the Workshop instead of directing specific methods.

As a result, since the inception of the AFIT-led Should Cost Workshop in March 2016, 31 IPTs have attended with 89 SCIs identified and $1.039 billion reported by the IPTs as potential savings over the upcoming 1–3 year timeframe.

Much has been learned from these workshops and AFIT is continually listening to its attendees. Lessons Learned include: Depth of analysis is highly dependent on data access and having the right technical and functional expertise on the program. Full team representation equated to more in-depth analysis while less than full team representation resulted in mostly summary or high level results. Additionally, cross-talk and idea sharing proved valuable to IPTs and often bridged the gap when teams were not fully prepared.

As a result, several recommendations are offered:

- Fully examine potential IPTs to determine if they are ready for the Workshop. Attendees must be post milestone A in order for significant data to exist to analyze.
- Contact IPT leadership to ensure their participation in Workshop kick-off and/or final day closeout.
- Group workshops on similar program type if possible. This is aimed at creating synergies between programs and increasing idea sharing.

Research Issue and Objective

The Air Force Should Cost Workshop is intended to produce multiple Should Cost Initiatives (SCIs) after the rigorous application of AFIT's established seven-step process by actual members of selected program IPTs. SCI data produced during a Workshop is not consistently documented to determine if actual SCI savings were achieved or even realized. Follow-up analysis will determine the status of previous SCIs and produce useful data to construct predictive models to identify trends, patterns, and commonalities of SCIs. Data gathering will produce numerous categories of SCIs that will be analyzed with tools such as Pareto Analysis.

Methodology Development / Background

Ashton Carter, Under Secretary of Defense, Acquisition, Technology, and Logistics (AT&L) issued his memorandum on June 28, 2010, titled *Better Buying Power: Mandate for Restoring Affordability and Productivity in Defense Spending*. Carter states his goal of "delivering better value to the taxpayer and improving the way the (Defense) Department does business." This was the start of Better Buying Power (BBP) 1.0 which introduced a new paradigm toward cost savings by adopting government practices that encourage efficiency through the use of Should Cost management (Carter, 2010, p. 1).The Air Force responded the following year by issuing further guidance from Jamie M. Morin, Assistant Secretary of the Air Force–Financial Management and Comptroller, dated June 15, 2011, entitled *Implementation of Will-Cost and Should Cost Management* (Morin, 2011, p. 1). This document challenged program managers to drive productivity improvements into their programs during contract negotiation and program execution by conducting Should Cost analysis.

Air Force acquisition leadership at SAF/AQX and Air Force Life Cycle Management Center (AFLCMC) understood that to fulfill the goals of BBP's emphasis on affordable programs and development of cost savings via Should Cost Initiatives, there were several obstacles faced by the Air Force acquisition workforce. Chief among them were educating a large, diverse workforce on Air Force Should Cost policy and a methodology or approach for implementing that policy, along with an ability to assist IPTs in following the methodology to discover SCIs for their specific programs. SAF/AQX had previously worked to provide education on the history of BBP and Should Cost, along with current policy, as part of a Should Cost Workshop developed and taught by the University of Tennessee under contract to SAF/AQX. The Workshop was delivered as a live, classroom-based course at a variety of Air Force bases, but budgets and time allowed for only a few offerings during the existing contract. A solution was needed that could more quickly impact a much larger percentage of the Air Force acquisition workforce. SAF/AQX enlisted the support of Should Cost subject matter experts at AFLCMC/AQP, working in conjunction with faculty at the School of Systems and Logistics at the Air Force Institute of Technology (AFIT/LS) to develop a two-phased approach to addressing these obstacles.

Phase 1 was to develop an asynchronous distance-learning course for the Air Force acquisition workforce to educate them on and provide wide exposure to the current Air Force policy on Should Cost. The course, entitled SYS 190 Air Force Should Cost Fundamentals, provided historical background on BBP and Should Cost, along with current DoD and Air Force policy on Should Cost and Should Cost Initiatives. AFIT faculty, working with SMEs from AFLCMC, developed a new, seven-step methodology for discovering and documenting Should Cost initiatives. SYS 190 describes the seven-step process, provides examples and illustrations of how the process could be applied, and also provides video-based case studies of a variety of Air Force programs that had discovered and implemented successful SCIs, with key team members or project managers from each project describing how their SCIs were discovered.

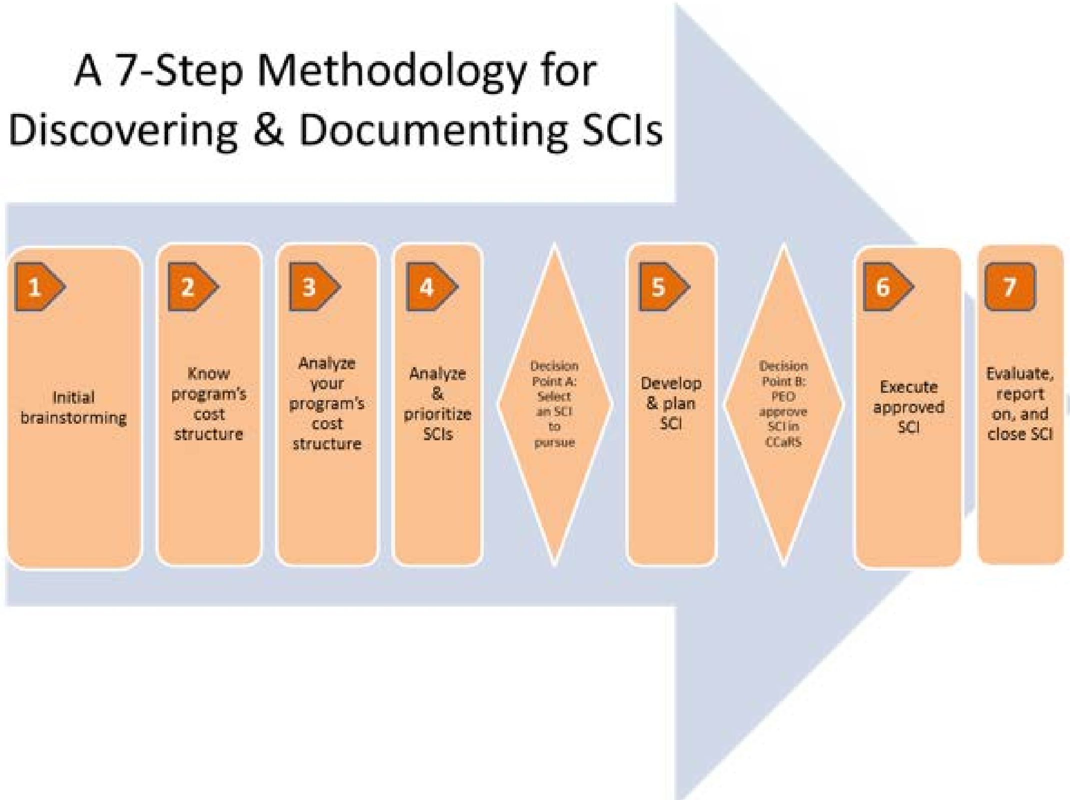


Figure 1. A 7-Step Methodology for Discovering and Documenting SCIs

The seven-step process as developed by AFIT is summarized as follows (see Figure 1):

Step 1 is to form a cross functional team and perform an initial brainstorming session. This is an opportunity to discover "low-hanging fruit," by encouraging creative thinking, while addressing aspects of a program that seem wasteful or inefficient. This is an attempt to capture some obvious ideas right up front before completing a deep dive into the program's cost structure looking for other SCIs.

Step 2 is to examine and know the program's cost structure. The goal is to understand where the majority of program costs lie, and to understand the key drivers of those costs and how those drivers might be influenced.

From that understanding of the program's cost structure, we turn to **Step 3**, where the teams attempt to take that detailed knowledge of the program's cost structure and analyze it from many different perspectives to find opportunities to save costs and/or increase efficiencies on the program. This is arguably the hardest and most time-consuming step in the process. By the time the teams reach **Step 4** in the process, functional teams have at least a few potential SCIs identified, both from the Step 1 brainstorming sessions and from the deep dive and analysis in Step 3. In Step 4, the list of potential SCIs is analyzed and prioritized emphasizing which SCIs to pursue first, second, and so on.

At the first decision point, Decision Point A, the team selects a particular SCI to pursue. The resultant SCI transitions to **Step 5**, where a Plan of Action is developed to implement the SCI and estimate the cost savings (in terms of cost avoidance and budget savings). It is also in this step where the initiative is loaded into the Comprehensive Cost and Requirements System (CCaRS) database. CCaRS is the official Air Force repository for this information, and provides a “dashboard” for leaders and decision makers to gather data for their needs while minimizing data collection requests from programs. It should be noted here that there is nothing prohibiting a program from planning and implementing multiple SCIs simultaneously.

Decision Point B is the point where the SCI is presented to the PEO for approval. **Step 6** begins after PEO approval, where work begins on the approved SCI using the plan developed in Step 5. Estimates and actuals for savings and realization dates are to be input and tracked using CCaRS.

Finally, **Step 7** is where the SCI is completed. The success (or failure) of initiatives are documented, along with details and lessons learned. If successful, the realized Should Cost savings are recorded, an update to the reinvestment recommendations is submitted (as required) for approval, and the SCI is closed.

These steps are straight-forward and easy to understand, but the work at each step is often detailed and difficult. AFIT believes there is nothing magical about this process, and others could come up with a slightly different process that works just as well. Whatever process a program follows, understand that just following the steps will not automatically produce SCIs. Finding SCIs is the result of detailed analysis and thinking by a diverse group of individuals who know the program—often difficult, time-consuming work. With this foundational knowledge available to the entire Air Force acquisition workforce, SAF/AQX and AFLCMC then asked AFIT to move to Phase 2, which was to develop a facilitated workshop where AFIT faculty SMEs would help guide complete IPTs to follow the seven-step process to discover SCIs on their projects.

Workshop Delivery

The Air Force Should Cost Workshop (WKSP 0656) is sponsored by SAF/AQ and is designed to be delivered to IPTs rather than individual students. AFIT is subscribing to the idea that an IPT populated by qualified functional departmental representatives is best for discovering and analyzing SCIs. Functional teams typically consist of a representative from program management, financial management, contracting, engineering, logistics, cost analyst, functional project experts (i.e., machinist, aircraft specialist, technology expert, etc.), and other subject matter experts (SMEs) that are needed to fully examine Should Cost Initiatives. This reduces the possibility of focusing too much on individual disciplines and not considering the full team’s experiences. The objective of the Workshop is for each IPT to produce at least one SCI to be presented to leadership for approval. Most IPTs develop multiple SCIs which broadens the likelihood of approval. AFIT instructors act as facilitators during the Workshop (rather than traditional course instructors) to guide IPTs through the process.

In order for participants to better understand the Workshop concepts, IPTs are required to take SYS 190 as a prerequisite before attending the live Workshop. AFIT facilitators explain the Air Force seven-step methodology (see Methodology Development / Background above) during the two-day Workshop. Steps 1–4 are presented in detail as these are the core processes needed for SCI development. Ample time is given for the IPTs to gather in breakout groups to go through each step; in fact, the majority of class time is spent in breakout sessions developing potential SCIs. This is where IPTs can dig-in and analyze their projects down to the individual cost element level. This level of detail forces the IPTs to (fundamentally) question everything. Facilitators encourage IPTs to grab the “low hanging fruit” or the obvious costs savings that only they know. This combined with Pareto Analysis produces ideas that can be analyzed to give the “biggest bang for the buck” for their programs. After each step, the IPTs gather together to out-brief their results to the other teams. This encourages one of the more valuable aspects of the Workshop—idea sharing across IPTs. Also called "cross-pollination," IPTs share their SCIs, ideas, and experiences. It is not unusual for other IPTs to gain new insight into processes that were not being considered before. Examples include the following:

- Using government testing facilities instead of paying for contractor-owned space
- Utilizing government furnished property (GFP) when applicable instead of purchasing the equipment by vendors
- Using simulators more extensively instead of actual flights, live fire tests, and other costly real-world testing
- Sharing base-wide contacts to help expedite activities.

This is where many of the "Should Cost non-believers" begin to see the value and importance of the Workshop. The second day of the Workshop ends with a discussion of steps 5–7. While these steps are not the focus of the Workshop, it is important for IPTs to understand and follow through with actually documenting, planning, approving, executing, and eventually closing each SCI.

Workshop Evaluation: Commonalities

The AFIT team is pleased with how the Workshop is being delivered based upon results (number and value of SCIs developed) and student feedback from Workshop surveys received after each offering. There appears to be a predictable pattern developing that determines how successful a particular Workshop will be given the enthusiasm expressed by potential IPTs. While all IPTs have been professional, engaged, and hardworking, some are more prepared than others. Common signs of less productive IPTs include lack of necessary data to perform needed analysis, lack of representation along functional lines, inexperienced or new team members assigned, poorly documented cost data, and teams not understanding their cost baselines. Common signs of more successful IPTs include fully represented functional teams with experienced members in attendance, one or two IPT members are in charge and direct their team's activities (they seem to work more efficiently), clear support from their leadership concerning the importance of Should Cost, complete access to costing data (and any data needed), and team members being open to new ideas. For additional insight, see the section entitled Workshop Lessons Learned below.

We have observed when similar IPTs are in attendance during a Workshop, additional benefits present themselves. The most obvious is a familiarity with systems and the overall mission of the programs. One example was when a Workshop had two munitions IPTs present. During post-step evaluations, questioning was more intense and detailed. Idea sharing was effective and highly productive. A camaraderie develops to actually encourage similar IPTs to try and help each other to root out costs and further develop SCIs. As mentioned earlier, this also produces new contacts that IPTs can reference and consult with in the future. One would assume that similar programs at the same location would be interacting but our team has noticed this is not a good assumption. Programs can be stove-piped and isolated for various reasons. Increased interaction between similar teams can have only positive effects.

Workshop Should Cost Data

Since its beginning in March 2016, the Should Cost Workshop has accumulated data across 20 Workshop offerings at eight Air Force installations. A total of 31 IPTs participated, producing 89 SCIs that were divided into 12 categories totaling $1,039,743,000 of potential Should Cost savings. Table 1 illustrates Workshop locations with associated IPTs, SCIs, and potential savings. Note: For the purposes of this paper, we are not disclosing the association between SCIs and their individual IPT due to implied confidentiality given during the Workshop. All data is accessible through existing reporting channels (see CCaRS discussion below) to those with proper access to the system.

Table 1. AFIT Should Cost Workshop Locations Data

Location	IPTs	SCI	Dollar Savings
Eglin AFB	3	11	$6,603,000
Gunter Annex, AL	3	6	$130,000
Hanscom AFB	5	27	$286,350,000
Hill AFB	7	2	$6,100,000
Robins AFB	2	7	$87,780,000
Tinker AFB	1	3	$48,250,000
WPAFB	10	33	$604,530,000
*Arlington, VA**	*1*	*1*	*$22,400,000,000*
	31	**89**	**$1,039,743,000**

*Arlington, VA IPT and SCI not included with $22.4B F-35 in Workshop totals. F-35 was an unusual SCI given the size of the program and 65-year amortization (see discussion below).

While we decided to not disclose specific individual IPT data, we would like to share a sampling of the collected SCIs. While not associating them with their particular IPT, a sampling by SCI type will help to illustrate the variety of SCIs developed by the IPTs. This data is illustrated in Table 2. Additionally, we are including a sampling of the IPTs that attended our Workshop in Table 3 divided into ACAT I (large) and ACAT II & III (small) projects to illustrate the diversity of the programs that attended.

Table 2. AFIT Workshop IPT SCI Examples by SCI Type

	SCIs	Type
1	Contractor Installs	Contractor
	Competing Support Equipment	Contractor
	Quality 2nd vendor	Contractor
	Buy Kits via Small Business	Contractor
	Contractor Travel	Contractor
	Production Strategy (LRIP) Procure 50/50 split w/ both offerors	Contracts
2	Implement Mil Cloud	Data
	Data Rights	Data
3	MIDS JTRS Lot Buy	Hardware
	ARMS CIE Tech Refresh	Hardware
	Modularization/refactorization	Hardware
	Use GFP instead of actual equipment purchase	Hardware

4	Remove plugin and reduce man-hours for PGMPs	Labor
	Decrease SEPM	Labor
	Reduce Security Guard Costs	Labor
	Reduce Training Costs & Schedule	Labor
	Manpower Rate Reduction	Labor
5	Engine Overhauls	Maintenance
	Heavy Maintenance Intervals	Maintenance
6	Installation Synergies	O&S
	Competing Communication Modifications	O&S
	Competing Interior Modifications	O&S
7	McAAP Process Improvement (TY$M)	Process
8	Limit changes to requirements once the build / fix is started	Requirements
	Tracking of Requirements	Requirements
9	Reduce scope of 520th sustainment to only BCSS maintenance	Scope
	Production Rephasing (TY$M)	Scope
10	Create central repository for baseline code	Software
	Incorporate Mil Cloud capability	Software
	Mirror systems at SSF and OITF	Software
11	Arena test: reduce arena test from 4 to 2 (1 for each contractor)	Test
	Captive flight test: eliminate captive flight test from plan based on SE analysis	Test
	Integrate Flight Test: skip DT and go straight to OT	Test
	Use government owned test facilities instead of contractor's	Test
	Use simulators in lieu of actual flights, live fire tests, and other real-world testing	Test
	Reduce Test Durations	Test

Table 3. Sample of AFIT Workshop Participating IPTs—By ACAT Level

Large IPTs – ACAT I
Airborne Warning and Control System (AWACS) Block 40/45
HH-60 Block 162
F-35 Joint Strike Fighter (JSF)
Presidential Aircraft Recapitalization
Targeting and Geospatial Intelligence (T&GI) Program
Global Hawk
F-15 Multifunctional Information Distribution System (MIDS) Joint Tactical Radio System (JTRS)
Smaller IPTs – ACAT II & III
Bomb Live Units (BLU-129/B)
Bomb Live Units (BLU-134)
DP / SP Suite Upgrade
Integrated Aircrew Ensemble (IAE) Program
Joint Direct Attack Munition (JDAM)
M-Code
UH-1N - AETC FLIR Replacement

While this paper focuses on the AFIT Should Cost Workshop, we note the existence of more than six years of Should Cost data that is currently available via CCaRS. CCaRS education is part of our Workshop (Step 5 and Decision Point B) where our team instructs students to enter in their SCI data for reporting, tracking, and eventual closure. CCaRS contains a wealth of information that is useful for comparison purposes and helps us understand how our Should-Cost Workshop fares against earlier offerings (non-AFIT).

Individual IPT Discussion: F-35 Joint Strike Fighter

In July 2016, the AFIT Should Cost team delivered a Workshop in Arlington, VA to the F-35 SPO to assist with their ongoing desire to reduce sustainment costs. The F-35 team was anxious to see if our methodology would help with their goal of reducing $300 billion from the program's sustainment budget. The System Program Office (SPO) already employed a robust Should Cost process that we were able to help improve upon. Our team knew this Workshop was going to be different when Step 1 of our methodology (Brainstorming) that usually takes two or three hours took the entire first day. The SPO produced 111 brainstorming ideas. Out of all these ideas, it came down to 12 that were seriously considered and only one that was presented. The team was able to convince their management to consider increasing simulator training time by 9% while reducing expensive in-flight training for carrier landings. This landing is considered riskier than others so the suggestion of more simulator training and less actual landings, in this case only, was acceptable. When calculated over the lifespan of the aircraft, it was determined to save $22.4 billion! The main concern with the identification of the $22.4 billion savings was that it was spread out over the life cycle of the aircraft which is currently estimated to be the year 2080 (extended from 2065) which makes any estimates uncertain in our minds.

Data Analysis

Our analysis began by looking at the limited aggregate data compiled from the 20 AFIT Should Cost Workshops offered from March 2016 to February 2017. Part of our objective was to determine if there are any SCI correlations between the offerings. Figure 2 looks at the SCI Pareto distribution between the 13 SCI types and the dollar amount associated with each. The SCI types that were derived from the IPTs are (in order of most dollars to least) maintenance (mostly aircraft), contractor support/service, operating and support (O&S; modifications and installations work), software, contracts, unclassified, hardware, scope, process, test, labor, requirements, and data (see details below).

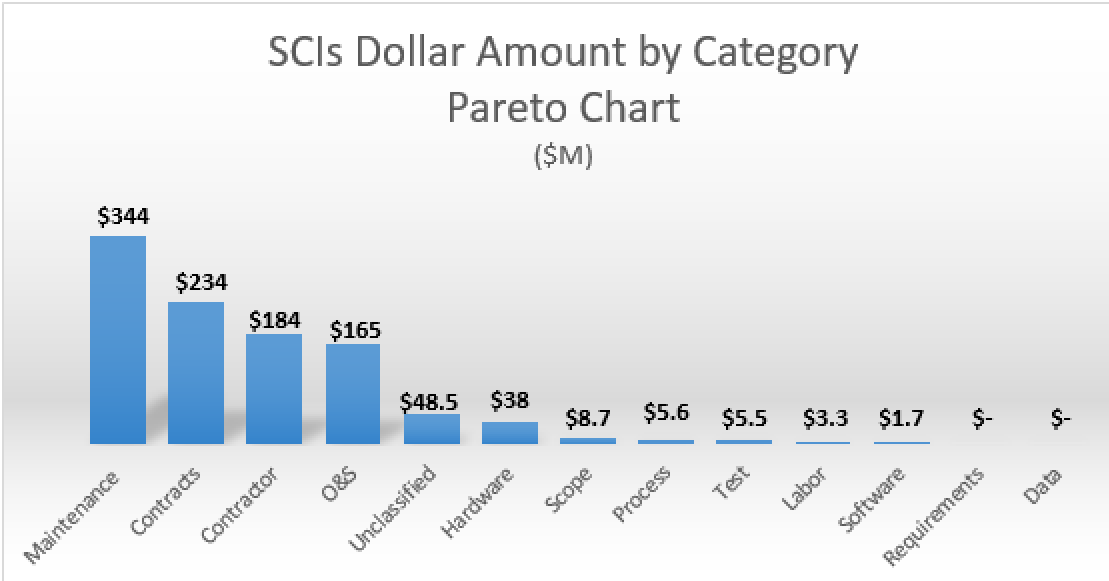


Figure 2. SCIs Dollar Amount by Category Pareto Chart

- Maintenance was by far the highest dollar amount which was associated with two SCIs that entailed engine overhauls and heavy equipment maintenance overhauls. These are examples of the occasional large dollar amount SCIs that were not the norm. This doesn't mean they are not important, rather these are outliers as compared to the majority. This IPT was excited to find such savings for their aircraft and was confident their leadership would approve, though no decision has been made as of the writing of this paper. This also includes reference to the F-35 IPT listed in Table 1 above. This IPT is not included in our dollar totals because of its ability to skew the dollar amounts beyond reasonable comparison. (See F-35 discussion above for more details).
- Contracts were a common target for cost savings, mainly through proposed purchase price reductions and cost sharing. Contracts (and the contracting effort in general) are targets for cost savings given the DoD tendency to rely on Firm Fixed Price (FFP) contracts. This is an area were IPTs could encourage more flexibility that would allow program specific circumstances to influence the selection of contract type more often then they currently do.
- Contractors Support/Service was next, and not surprising given the amount used by the Air Force (and DoD). Though once again 45% of the savings on the contractor SCIs were due to one large contractor installation effort. This presents a limitation based on our small sample size but it does not diminish the effects contractors have on Air Force programs. There is a common belief that contractor costs can be better managed, resulting in cost savings.
- Operating and support (O&S) represents costs associated with modifying, maintaining, installing, and supporting equipment and were big items considered from our aircraft IPTs.

- The unclassified category pertains to those SCIs that were assigned a dollar value but were not detailed enough by the IPTs. The AFIT team was advised that the IPT may get back to us with the details at a later date.
- Hardware was the result of only two IPTs that were updating system hardware. The AFIT Should Cost team thought this would be more of a common savings area, but IPTs explained many upgrades are in the form of Commercial Off-the-Shelf (COTS) hardware with little bargaining room due to smaller lot buys over many years. Suggestions were made to pursue lifetime buys, though the effectiveness of this depends on the program.
- Items such as scope, process, test, and labor tended to be low dollar, but representative of a large number of SCIs.
    - Labor can represent a large portion of program cost (especially service-oriented contracts) where significant savings can be realized. It can be noted that IPTs working with large integrators like Boeing, Lockheed Martin, Raytheon, etc. are frustrated with DCMA and other defense organizations that issue pre-negotiated wrap rates that cannot be changed. IPTs do concede they may not know the big picture when it comes to large vendors but the ability to negotiate more on the program/project level would be welcome.
- Requirements and Data received a total of six SCIs but no dollars associated. Requirements (only one program) listed dollars as TBD due to pre-milestone B program status and would supply once finalized. Data focused on the data rights of one program which were still under negotiation with no dollar amount available.

The AFIT Should Cost team is satisfied at the number of SCI types which we interpret as IPTs performing deep analysis and being creative in generating cost saving ideas. The IPTs realize that not all of their ideas will be successful (or even approved) but this did not inhibit their desire to look deeper for savings.

The next area of research includes the actual number of SCIs developed, as illustrated in Figure 3. While one would surmise that high dollar amounts equate to high SCI count, they would nevertheless be mistaken. Some of the highest SCI counts relate to the lowest dollar amounts and vice versa. Two of the smallest dollar amounts, Test and Labor specifically, account for over 22 of the 89 SCIs (or almost 25%). This apparent randomness can best be observed in Figure 4 where we transposed Figures 2 and 3 to get the combined view. It would appear that there is no correlation between the number of SCIs and anticipated savings to a project. Our takeaway from this is not to underestimate what any IPT can achieve when examined in a non-attributive setting and allowed to challenge the status quo. This does not guarantee approval, but it does broaden the awareness of individuals and provide a catalyst for critical thinking focused on ways to save and be more efficient within their IPTs.

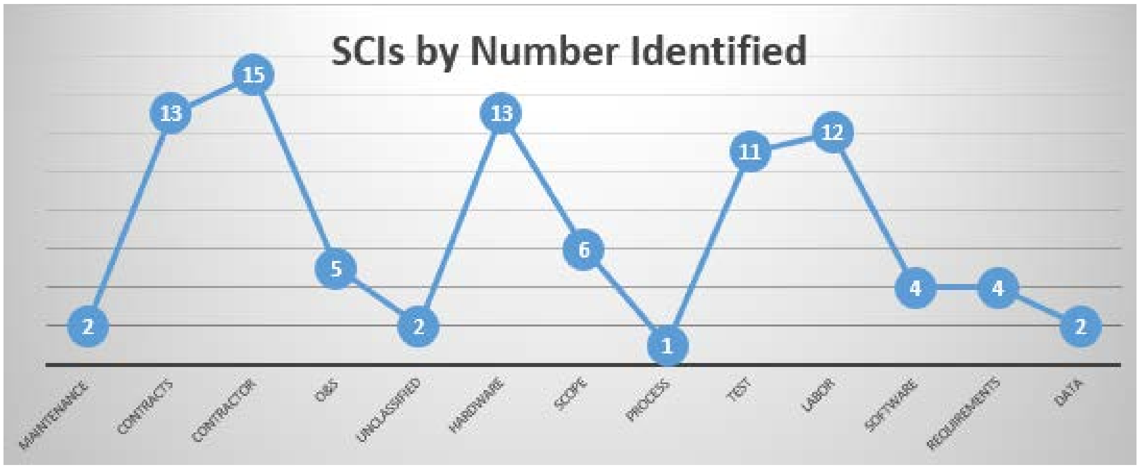


Figure 3. SCIs by Number Identified

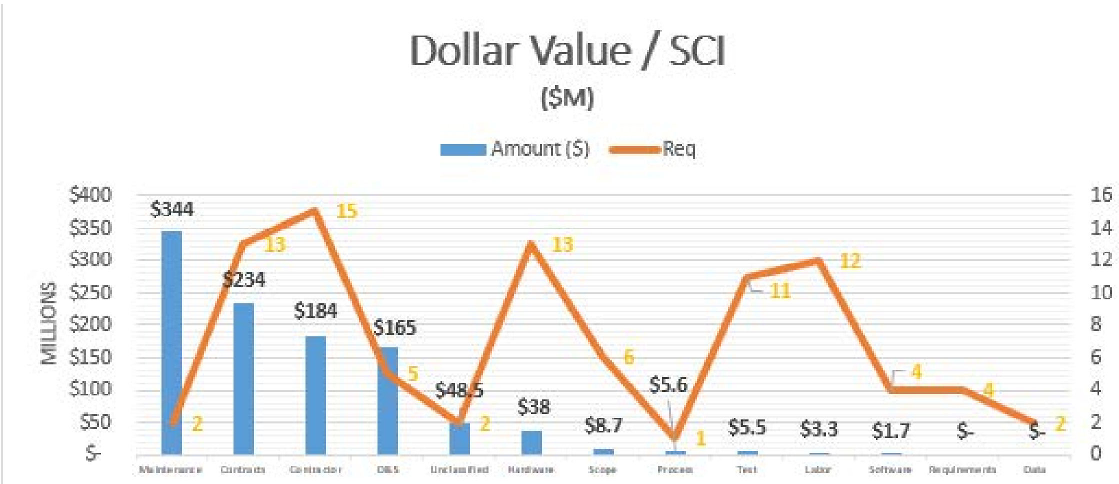


Figure 4. Dollar Value per SCI

The AFIT Should Cost team analyzed more than just the data accumulated from the Workshops. General behavior played just as important a role when considering how it can affect IPT performance and ultimately Workshop results. Workshop Delivery (see section above) talked about the mechanics of the Workshop but not the interactions between the IPTs and individual team members. As a result, we have compiled a listing of Workshop lessons learned from the AFIT Workshops that incorporate our observations from the IPTs. No one particular lesson learned was unique to any one IPT; rather, these lessons learned are considered trends seen throughout the Workshops.

Workshop Lessons Learned

- Coordination with local IPTs is vital for success. Proper screening of candidate IPTs helps eliminate teams that are either too early in the acquisition life cycle to be effective (pre-milestone A) or too close to project closeout.
- There is a tendency for participants to be distracted by current work assignments, with many either missing large parts of the Workshop or not returning at all. PEO/PM involvement is required to mitigate this issue.

- IPT/base leadership support is key to success. Workshops that included leadership endorsement either in person or via IPT communication exceeded expectations. Where this support was lacking, Workshop IPTs tended to have missing or incomplete data, non-optimal program team representation (i.e., not all functional areas represented), inexperienced participants, and a higher propensity for attendees to show lower interest.
- Depth of analysis is highly dependent on data access and having the right technical and functional expertise on the program. Full team representation equated to more in-depth analysis while less than full team representation resulted in mostly summary or high level results.
- Cross-talk and idea sharing proved valuable to IPTs and often bridged the gap when teams were not fully prepared.
- IPTs frequently request a listing of existing SCIs from past Workshops as a guide to discovering their own SCIs. While examples do exist, the AFIT facilitators are reluctant to provide this information because we want the IPTs to go through the process and discover their own SCIs rather than selecting from a pre-existing list. A compromise may be reached where a listing of program specific SCIs is provided after Step 4. This would be a reference list that may be added if deemed useful. Another option would be to offer SCIs from unrelated programs (i.e., show munition SCIs to an aircraft IPT) where team members can get the idea of the level of detail / focus, which they can apply to their analysis. AFIT is still analyzing the benefits of each technique.
- Whether prepared with data or not, most participants said they were better informed and more prepared to engage in Should Cost Initiatives after the Workshop, than before. Most of the “reluctant participants” became valuable contributors.

Every course or workshop developed by AFIT has a goal of educating and informing attendees with new or better means and methods to help foster professional success. This is certainly true for the AFIT Should Cost Workshop. However, AFIT does have limitations when it comes to influencing student behavior once they leave the classroom. Steps 5–7 of our methodology instructs IPTs to enter their SCIs into CCaRS so they can be documented, monitored, and tracked. This is our primary way to measure how successful SCIs are for each program. Unfortunately, this is the area that could use the most improvement. Should Cost data in CCaRS (limiting the search from March 2016 to February 2017) was analyzed to determine how many IPTs had actually entered data within the time period AFIT provided Workshops. Between March 2016 and February 2017, 108 IPTs DoD-wide documented 228 SCIs. Of these totals, only six IPTs actually recorded SCIs in CCaRS; the number of SCIs recorded by these six IPTs in CCaRS was eight. This means that 5.5% of the IPTs entered their data (6 out of 108 IPTs) and 3.5% of the SCIs were from AFIT led Workshops (8 out of 228). To be fair, it is unknown how many DoD-wide IPTs did not enter their data because CCaRS does not provide such information. Regardless, the AFIT team would like to see these numbers improve. The cause of these low numbers can be interpreted in two ways: (1) IPTs are continuing to analyze their SCIs and waiting to attain leadership approval prior to entering data into CCaRS, or (2) once the Workshop ends, IPTs quickly lose focus on SCIs and return to their normal duties. Given the amount of time that has passed since beginning the Workshops, it appears that option 2 is the most likely candidate. With increases in PEO/IPT leadership involvement, the AFIT team believes IPTs would be more motivated to complete the Should Cost methodology by fully documenting SCIs in CCaRS.

Our team will continue to encourage PEO/IPT leadership to become more engaged by communicating the benefits of Should Cost analysis.

Recommendations

There are at least as many means and methods for improvement as there are program types and procurements. In an effort to continue advancing should cost efforts to produce those means and methods, and provide participants with better potential to produce initiatives, the AFIT workshop facilitators and other faculty have provided the following items as recommendations for consideration to improve workshop results. These recommendations are the result of actual Air Force Should Cost Workshop observations and experiences with the IPTs involved:

1. Fully vet the potential IPTs to determine if they are ready for the Workshop. Attendees must be post milestone A in order for significant data to exist to analyze.
2. Contact IPT leadership to ensure their participation in Workshop kick-off and/or final day closeout.
3. Present sample SCIs during Workshop facilitation to aid IPTs in their individual SCI development. These can be grouped by program type, size, or sustainment level if desired. Sample SCIs can also be provided based on unrelated programs. The AFIT team does not want to simply provide a list of SCIs that can be cherry picked by the IPTs. Our goal is for individual IPTs to perform their own analysis. Presenting SCIs from an unrelated program (i.e., an aircraft SCI for a munitions program) will give IPTs an idea of what is expected.
4. Group Workshops on similar program type if possible. This recommendation is aimed at creating synergies between programs and increasing idea sharing. There is still value in having diverse program types attend a Workshop together, so we don't recommend strict segregation by program type.
5. Off-site facilities are preferred over base-provided training locations. When away from the base, IPTs appear to focus better and are not as tempted to return to the office where they can become distracted or potentially not return to the Workshop.
6. During the Brainstorming session in Step 1, AFIT facilitators need to be continually checking on the IPTs to ensure the teams are staying on topic. The tendency to go off on unrelated tangents exists and can be tolerated from time to time but should not dominate the session.
7. Frequent student feedback suggests spending more time on CCaRS for data entry. Usually this is performed by the IPT program or financial managers, but other team members are showing interest so additional time should be spent going into more detail of the application.

The recommendations we provide above are some common themes we have found that foster successful Should Cost Workshop outcomes. However, there is no single silver bullet we can point to that ensures success in these Should Cost Workshops. Rather, it is a combination of factors, some under the control of the program IPT, and some under the control of the Workshop facilitators, which enable program IPTs to successfully identify realistic SCIs. Paying attention to these factors is certainly helpful, but we want to stress that finding realistic, actionable SCIs with the potential for generating savings requires conscientious preparation and hard work on the part of both the program IPT and the workshop facilitators. We suggest that program IPTs use these recommendations as a starting point for planning their own program's journey through the trials, tribulations, and ultimate rewards to be reaped by applying Should Cost Management to their programs.

References

Carter, A. (2010, June). *Better buying power: Mandate for restoring affordability and productivity in defense spending* [Memorandum]. Washington, DC: DoD.

Morin, J. M. (2011, June). *Implementation of will-cost and should-cost management* [Memorandum]. Washington, DC: Department of the Air Force.

Creating a "Should Cost" Culture Through Opportunity Management

David Riel—is currently Professor of Acquisition Management for DAU's Midwest region with Level III certifications in PM and PQM. His 20-year USAF career includes working on a variety of programs, including the F-22 Advanced Tactical Fighter and the Global Hawk Unmanned Aerial System (UAS). Riel also spent several years working as Senior Manager of Manufacturing and Continuous Improvement with one of our defense industry partners. [david.riel@dau.mil]

Abstract

The 2016 *Performance of the Defense Acquisition System Annual Report* states, "The institution of 'should cost' management … has been a success and should be a permanent feature of the DoD's acquisition culture." Yet, inculcation into culture implies a mindset that executes cost saving opportunities throughout all levels of the acquisition workforce to the lowest level workers. Initiatives, such as Configuration Steering Boards and USAF's "Bending the Cost Curve" are important and communicate leadership engagement; however, creating a "should-cost" culture requires every level of the acquiring organization to continuously seek both small and large cost saving ideas. One could posit that the system engineering community has transformed from merely managing risks to expanding into exploring opportunities. However, the question becomes, are opportunities being as aggressively pursued and managed as risks? By conducting surveying, this research studies the perceptions of acquisition professionals as to the implementation of opportunity management for the purpose of determining the depth that "should-cost" philosophy has penetrated organizations. This study discusses methods to drive the "should-cost" philosophy deeper into organizations, along with the role that defense acquisition education could play.

Introduction

Although the concept of "should-cost" has been around since shortly after World War II as both technical and organizational complexity soared, the recent rendition of should-cost is traced to the Better Buying Power initiative, established by Ashton Carter, then Under Secretary of Defense, Acquisition, Technology, and Logistics (USD[AT&L]) in 2010 (Burt, 1972; Husband, 2014). Recently, the Department of Defense (DoD) has touted to Congress that "lower contract costs, reduced cost overruns, and arrested cost growth" can be directly tied to BBP's "should cost" initiative (Maucione, 2017). Accordingly, in the 2016 *Performance of the Defense Acquisition System Annual Report*, the authors state that "the institution of 'should cost' management and its consistent emphasis over the last 6 years by the acquisition chain-of-command has been a success and should be a permanent feature of the DoD's acquisition culture" (pp. xviii, xix). This statement is not surprising as Carter, in a 2011 *Defense AT&L Magazine* article, co-written by John Mueller, wrote that "There are no silver bullets; each PM must find solutions that fit his or her specific program. In the final analysis, embracing the 'should cost' management paradigm represents a cultural change, not just a one-time event" (p. 17), along with "It is not a one-time fix but a change in the culture of our government teams and our contractors" (Carter & Mueller, 2011, p. 18). Yet, inculcation into culture implies a mindset that executes cost saving opportunities throughout all levels of the acquisition workforce down to the lowest level workers. Yet, as Husband (2014) points out, "Unlike industry, which is driven by profits, government PMs often focus solely on risks and pay insufficient attention to cost-reduction opportunities" (p. 571). DoD Director of Defense Pricing, Shay Assad, said that instilling the acceptance of should-cost was a great struggle for the DoD and that the bureaucratic mindset, that is, culture, needed to start thinking differently from historical-based pricing, that is, "will cost" to "should-cost" (Maucione, 2017; Risor, 2011). Also, inculcation into culture implies that values and artifacts align with the pursuit of should-cost, which takes time and leadership attention (Hatch & Cunliffe, 2013).

For major programs, DoD 5000.02 requires that

> program managers, in consultation with the PEO, and the requirements sponsor, will, on at least an annual basis, identify and propose to the CSB (Configuration Steering Board) a set of recommended requirements changes to include descoping options that reduce program cost and/or moderate requirements and changes needed to respond to any threat developments.

While the program managers' efforts in bringing tradeoffs before the CSB is also a step in the right direction, reducing costly design decisions by capitalizing on the effects of many small changes will be difficult to achieve in a once yearly forum. For these types of tradeoffs, systematic, process-driven pursuit of small changes is required via other ways, for examlpe, opportunity management (OM).

Similarly, in 2015, the Air Force announced its "Bending the Cost Curve" initiative, featuring a cost-capability analysis program, which also offers cost-savings opportunities. However, this initiative only promotes the conversation during the pre-EMD contract award stage. Unfortunately, it is only during detailed design that many cost-capability opportunities become apparent, reinforcing the need for a continuous opportunity-driven mindset. All of these initiatives are important and communicate leadership engagement; however, creating a "should-cost" culture requires acquisition professionals at all levels of the organization to continuously seek and implement both small and large cost saving ideas.

With the replacement of the August 2006 R*isk Management Guide for DoD Acquisition* with the broadened June 2015, and subsequent 2017, *Department of Defense Risk, Issue, and Opportunity Management Guide for Defense Acquisition Programs* (ROI Guide; DoD, 2017), one could posit that the system engineering community has transformed from merely managing risks, both potential and realized (issues) to expanding into exploring opportunities. At the lower levels of the organization, opportunities, as defined by the ROI Guide, "are potential future benefits to the program's cost, schedule, and/or performance baseline, usually achieved through proactive steps that include allocation of resources" (p. 43), and that "risk and opportunity management support Better Buying Power initiatives to achieve should-cost objectives" (DoD, 2017, p. 43). Figure 1 from the RIO Guide highlights the relationship between should cost and OM, which elucidates the importance of OM at lower levels of an organization in the effort to create a "should-cost" culture.

Figure 4-1. Opportunities Help Deliver Should-Cost Objectives

Figure 1. Opportunities Help Deliver Should-Cost Objectives
(DoD, 2017, Figure 4-1)

However, the question becomes, are opportunities as well understood and are they being as aggressively pursued and managed as risk and issues? By conducting quantitatively-based, statistically-significant surveying, this research studied the perceptions of acquisition professionals as to the implementation of OM at the lower levels of acquisition organizations for the purpose of determining the depth that the "should-cost" philosophy via the lower level OM has penetrated their organizations. Using this data, the study posits methods for increasing the ability for leaders to drive "should-cost" philosophy deeper into their organizations and the role that defense acquisition education should play for the purpose of creating and sustaining a "should-cost" culture.

Should Cost

Although the concept of "should cost" is tied to the 1960s practice of employing a team of experts to complete a comprehensive, meticulous analysis at a defense industry contractor's facility, and limited as "what a defense system *ought to cost*, assuming reasonable economy and efficiency in the contractor's operations" (Burt, 1972, p. 3), Ashton Carter, in his BBP initiative, co-opted the phrase and broadened it to include challenging "the business-as-usual approach, with its underlying assumption that program costs will grow to match (or exceed) the independent cost estimate" (Carter & Mueller, 2011, p. 14). Carter and Mueller (2011) warn not to confuse the 2010 version of "should-cost" with the DFARS "should cost review" and its emphasis on production. The ultimate goal of the earlier should-cost review was to provide the government team with in-depth information in support of a better negotiating position and "set realistic objectives for negotiating the immediate contract" (Burt, 1972; Morin & Van Buren, 2011, p. 1). Instead, the "should-cost" initiative found in BBP demands across-the-lifecycle implementation, with a particular emphasis on "up-front planning and exploring engineering trades," as program managers "drive leanness into their programs by establishing Should-Cost estimates at major milestone decisions," by looking at every "cost element, including government costs, acquisition strategies, and any technique that could provide net savings" (Carter & Mueller, 2011, p. 16; Husband, 2014, p. 568; Morin & Van Buren, 2011, p. 1).

Congress, without using the term "should cost," also addressed the use of historical data for budgeting purposes when the Ike Skelton National Defense Authorization Act for Fiscal Year 2011 passed employing the following verbiage:

> (a) cost estimates developed for baseline descriptions and other program purposes … are not to be used for the purpose of contract negotiations or the obligation of funds; (b) cost analyses and targets developed for the purpose of contract negotiations and the obligation of funds are based on the government's reasonable expectation of successful contract performance in accordance with the contractor's proposal and previous experience. (Husband, 2014)

Willen and Garber (2011) posit that when implemented systematically, should cost can "reduce system costs by 5 to 15 percent and subsystems by up to 40 percent." In fact, William LaPlante, when Assistant Secretary of the Air Force for Acquisition (SAF/AQ), claimed over $2 billion in actual program cost cutting in recent years with more to come (Serbu, 2015). Experience has shown that benefits are best when a cross-functional team, employing various skills, is required to reap the should-cost review benefits (Bioto et al., 2012).

Organizational Culture Change Challenges

At its heart, "should cost" is about challenging the status quo, that is, the historical costs that form the basis of what a program is budgeted to. With the requirement of managing to should cost being a "long-term endeavor" (p. 17) and the inevitable resistance to change from the status quo, inculcation of should cost into the DoD culture will take considerable, persistent effort (Carter & Mueller, 2011, p. 14). Two major resistor elements are the DoD's mechanistic organizational structure, which tends to preserve the status quo, and the program managers' fear that "should cost" was just one more opportunity for higher headquarters to cut their funding (Husband, 2014). David Van Buren (Air Force Service Acquisition Executive) and Jamie Morin (Assistant Secretary of the Air Force [Financial Management and Comptroller]) in a joint memorandum confirm this trust issue by stating, "We recognize program managers have concerns about providing estimates that are lower than the budget, since DoD culture tends to use programming and budgeting to incentivize achievement"; however, they affirm that "this is not the intent of this [Should Cost] initiative" (Morin & Van Buren, 2011, p. 1.). Also, Carter sought to alleviate the fear of lost funding with his joint memorandum with the Under Secretary of Defense (Comptroller/Chief Financial Officer) Robert Hale, writing that after validation of actual savings by the Service Assistant Secretary (Financial Management and Comptroller), that "Savings would then generally be retained by the Service and reallocated to the highest priority needs as determined by the Service Secretary or a senior leader as designated by the Service Secretary" (Carter & Hale, 2011, p. 1).

The DoD's mechanistic organizational structure may present an even more challenging hindrance to culture change. Mechanistic organizations are depicted by elevated levels of hierarchical structure and hegemony; distinctly delineated roles and responsibilities; written policies and practices; specialized, standardized tasks; and centralized decision-making procedures, which research has shown to be restrictive to the innovation, flexibility, and creativity needed to identify opportunities (Hatch & Cunliffe, 2013). Mechanistic organizations are intended to achieve certain goals using fixed regulations, policies, events, or standards, which can be difficult in a complex, changing world ("Mechanistic Organizations," n.d.; Morgan, 1986). However, the advantages of a mechanistic organization are that formalization and control can lead to greater efficiencies through reduced variation and better predictability, which can help develop the institution of should-cost and OM, but will require added effort, as the mechanistic structure tends to hold fast to past practices, such as risk management (RM) only, rather than expand to include new practices, such as aggressively seeking OM (Fiol & Lyles, 1985; "Mechanistic

Organization," n.d). Organizational structure induces and is interwoven with the culture, including a confirmed "negative correlation between centralization and innovation" (Whittinghill, 2011, p. 17). "Should cost," and by extension OM, are a result of innovative behaviors and questioning status quo. For this to become cultural, strong leadership must encourage the creative thought needed to derive these opportunities.

Schein's Theory of Culture theorizes that a collection of basic assumptions, taken for granted by the culture's member, form the basis of a culture (Hatch & Cunliffe, 2013). For example, in America, the freedom of speech and religion, along with our democratic elections, are basic cultural assumptions. Even as our mindfulness of these basic assumptions diminish, they guide our perceptions, thoughts, and feelings (Hatch & Cunliffe, 2013). Schein considers values as the next level of culture, sharing that they provide the "social principles, goals, and standards that cultural members believe have intrinsic worth" (p. 169), leading members in their concept of right and wrong and leading to defined behavioral standards and expectations (Hatch & Cunliffe, 2013). The third and final level of Schein's hypothesis is artifacts, which are the outward indicator of values in the form of objects, verbal expressions, and behaviors (Hatch & Cunliffe, 2013). The U.S. Marines phrase *semper fidelis*, Latin for "always faithful," offers a superb case of an artifact derived from the values of honor and duty to country and fellow marines.

While the inherent control that comes with a mechanistic organization makes expressing new ideas, questioning performance requirements, and change difficult, one potential way of increasing the likelihood of a shift towards a "should-cost" culture is by increasing the emphasis on a lower-level, grassroots approach to cost savings, OM (Morgan, 2006).

Opportunity Management

OM can be described as an extension of a disciplined system engineering approach, and complement to the more well-known and better implemented RM (DoD, 2017; Pridgen et al., 2012). In fact, risk and opportunities can be seen as opposite sides of the same coin (Dester & Blockley, 2003, p. 83). Opportunities are defined in the 2017 RIO Guide, as the "potential future benefits to the program's cost, schedule, and/or performance baseline, usually achieved through reallocation of resources" (p. 43). OM is described by the RIO Guide as a support to "Better Buying Power initiatives to achieve should-cost objectives" (DoD, 2017, p. 43), as a way to "help offset cost or schedule impacts from realized risks" (p. 43). The OM process is similar to the RM process with both employing a five-step approach. Figure 2, taken from the ROI Guide, describes the OM process.

Figure 2. Opportunity Management Process
(DoD, 2017, Figure 4-2)

Programs are advised that they may either use the RM board or establish a separate OM Board (OMB), and upon disposition, track the opportunity via an opportunity register, which is analogous to the risk register (DoD, 2017). The RIO Guide describes four possible dispositions after the opportunity candidate has been evaluated for potential cost, schedule, and performance benefits, along with any additional risks potentially introduced. Options included are the following:

- "Pursue now – Fund and implement a plan to realize the opportunity. (Determination of whether to pursue the opportunity will include evaluation of the return of any investment when the opportunity would be realized, the cost, additional resources required, risk, and time to capture.)
- Defer – Pursue/cut-in later; for example, request funds for the next budget and request the S&T community mature the concept.
- Reevaluate – Continuously evaluate the opportunity for changes in circumstances.
- Reject – Intentionally ignore an opportunity because of cost, technical readiness, resources, schedule burden, and/or low probability of successful capture" (DoD, 2017, p. 45)

The RIO Guide advocates using any realized savings as an offset for any issues: yet, it also introduces the option to use the OM process to pursue more capability. While giving the warfighter added capabilities is a worthy goal, the better option may be to be more aggressive on "should-cost" goals. One of the potential problems with OM used for capability enhancement, as outlined by Conrow and Charette (2008) in their *Defense AT&L Magazine* article, "Opportunity Management: Be Careful What You Ask For," is that "unless tightly controlled, OM may exacerbate the enduring problem of requirements creep that plagues programs today" (p. 16), defining opportunities as "the potentially desired better- (or greater-) than-expected outcome of an event or situation that requires an additional allocation or reallocation of resources to pursue" (p. 16). One could argue that this criticism, while somewhat valid, is narrowly-focused on one potential use of OM. Quite the opposite application could be pursued using OM methodology, where cost-driving performance requirements could be challenged with the opportunity to trade off certain thresholds if cost saving and/or avoidance amounts warranted, with consent of the budgeting command, who determine overall affordability in accordance with DoD Instruction (DoDI) 5000.02, *Operation of the Defense Acquisition System*. In fact, the RIO Guide, while sharing the expectation that "high-return opportunities to improve the program life cycle cost, schedule and performance baselines” (DoD, 2017, p. 47) be evaluated and actively pursued, expectations are also for programs to “establish opportunity likelihood and benefit criteria in line with program ‘should-cost’ objectives” (p. 47). Analysis of the current DoD RIO Guide and defense acquisition education reveals the strong cultural preference for identifying and averting risks over identifying and pursuing opportunities. As described above, only five pages are dedicated to describing and instructing on OM within the 96-page RIO Guide. Also, an analysis of learning objectives of Defense Acquisition University (DAU) Defense Acquisition Workforce Improvement Act (DAWIA) curriculum demonstrates a strong preference for RM instruction over OM.

Assessment of Defense Acquisition University Learning Objectives

An analysis of the DAU curriculum focused on core DAWIA required courses for certification requirements revealed only one major terminal/enabling learning objective (ENG 301) for the topic of OM related to should cost, compared to the 22 major terminal/enabling learning objectives on RM. (Note that ENG201: Applied Systems Engineering in Defense Acquisition, Part I, is under development, and course objectives were not available.) Figures 3–16 show the results of a survey of DAU course objectives across the acquisition, engineering, contracting, and production, quality, and manufacturing learning objectives, where one might expect OM and RM objectives to most reasonably be found.

1	**Recognize the key drivers of the Department of Defense's Acquisition Management System.**
	Define Systems Acquisition Management.
	Recognize how risk (cost/schedule/performance) is at the core of acquisition management.
	Name the principle regulations governing defense acquisition and procurement.
	Recognize the requirement for effective safety and health programs for every defense acquisition program and top-level legislation that applies to our environmental concern.
	Identify the three major DoD decision-making support systems in defense acquisition programs.

Figure 3. ACQ 101—Fundamentals of Systems Acquisition Management Course Objectives

2	**Apply the risk management process as a basis for making sound acquisition program decisions**
	Relate the key tenants of IPPD to planning and executing an acquisition program
	Identify the barriers to successful IPT implementation
	Identify key acquisition best practices, including commercial practices that impact the relationship between government and industry.
	Identify the information required for a decision review and recognize the significance of the Acquisition Program Baseline, Key Performance Parameters, and Acquisition Strategy
	Identify the advantages and disadvantages of international armament cooperative development in an acquisition strategy.
	Identify long term supportability and sustainment strategies through the application of Product Support Business model (PSBM) and the 12 - Step Product Support Strategy Process Model.
	Capture the Product Support Strategy and specific planning execution details in the LCSP.
	Identify the five activities of the risk management process model.
	Use the risk management process to identify the major areas/sources of risk in an acquisition program strategy.
Use Technical Performance Measures to track progress in program risk areas during systems development.	
Identify the role of Early Operational Assessment (EOA) in reducing program risk.	
Relate the Acquisition Program Baseline (APB) to planning, control, and risk management in attaining cost, schedule and performance goals.	
Recognize the relationship between risk management and exit criteria.	
5	**Determine the role of contracting in the acquisition process and the major contractual contributions toward managing program risk.**
Apply the risk management process as a basis for making sound acquisition program decisions	

Figure 4. ACQ 202—Intermediate Systems Acquisition, Part A Course Objectives

9	**Analyze actual verses planned technical performance data in risk areas to indicate potential problems that may prevent a system from being operationally effective and suitable.**
	Identify potential risk areas based on technical performance data
	Identify the role of technical performance measures in the systems engineering process.

Figure 5. ACQ 203—Intermediate Systems Acquisition, Part B Course Objectives

3	**The student, while serving in the role of a contracting leader, will deteremine risk mitigation techniques to use in a contracting related problem.**
	Identify the reasons for the presence of risk
	Analyze risk handling techiques as stated in the DoD Risk, Issues and Opportunities Guide, June 2015 in a contracting issue
Identify industry risk factors during the contracting process	
Using knowledge of industry risk and personal perspective, formulate viable solutions to a contracting dilemma	
Using knowledge of industry risk and personal perspective,evaluate viable solutions to a contracting dilemma in order to reach a defensible decision	
19	**Given a program scenario and cost reduction opportunities, the student will develop engineering inputs for a should-cost management plan in accordance with in accordance with DoDI 5000 and the Defense Acquisition Guidebook (DAG).**
	Recognize the relationship between Affordability, Will-Cost, and Should-Cost baselines.
	Compare Value Engineering to should cost management.
	Describe an approach to implementing should cost management on an acquisition program.
	Given a set of executable cost reduction initiatives, analyze the initiatives from an engineering perspective.
	Given a set of executable cost reduction initiatives and set of constraints, develop an associated program should cost baseline.
	Given a program scenario, develop a list cost reduction initiatives to support should cost management in the operations and sustainment phase.

Figure 6. CON 360—Contracting for Decision Makers Course Objectives

17	**Describe the function of the Risk Management process as part of the Systems Engineering Process**
	Describe the purpose, inputs and outputs, and activities of the Risk Management Process
	Describe how Risk Management is used
	Outline the importance of risk identification
	Summarize risk mitigation techniques
	Explain the role of risk tracking

Figure 7. ENG 301—Leadership in Engineering Defense Systems Course Objectives

8	**Given DoD technical risk management problem solving scenarios, provide rationale for the selection and defense of a best solution using the guidance provided in the DoD 5000 series documents, DAG, and DoD Risk, Issue, and Opportunity Management Guide for Defense Acquisition Programs.**
	Given a risk management scenario, demonstrate the use of the DoD Risk Management process steps.
	Given a risk management and problem solving scenario, demonstrate use of the Risk Reporting Matrix using the Likelihood and Consequence Tables discussed in this lesson.
	Given a risk management and problem solving scenario (issue), defend the best option to pursue to control risk.
	Given a risk management and problem solving scenario (issue), determine specific implications of potential Design Considerations.

Figure 8. ENG 101—Fundamentals of Systems Engineering Course Objectives

5	**Given a capability development scenario, the student will analyze a capability need and develop a systems engineering risk assessment to support early acquisition life-cycle development planning in accordance with policy established in DoDI 5000.02 and the Defense Acquisition Guidebook.**
	Identify the systems engineering inputs to a Material Development Decision (MDD)
	Identify the major elements of an Analysis of Alternatives as described in the DAG.
	Identify sources for investigating technology opportunities and risks related to development of a needed capability.
	Derive a list interfacing systems and stakeholder organizations that will influence the requirements and constraints for a preferred materiel solution.
	Analyze an Initial Capabilities Document (ICD) to support development of system-level performance requirements.
	Assess a proposed materiel solution for technology insertion risks and opportunities.
	Assess user capability requirements and the related preferred materiel solution for risks related to systems engineering design considerations.

Figure 9. ENG 202—Applied Systems Engineering in Defense Acquisition, Part II Course Objectives

10	**Given a system development scenario and system architectural description, the student will develop appropriate system security risk mitigations (system security features), in accordance with the Defense Acquisition Guidebook and concepts discussed in class.**
	Analyze a system's CONOPS and architecture description for critical system functions and related critical program information.
	Analyze a system's CONOPS and architecture description to identify potential System Security risks.
	Evaluate a system CONOPS for tradeoffs between interoperability related requirements and system security requirements.
	Given a set of system security risks, develop a mitigation strategy to reduce the risks.
11	**Given a system development scenario, the student will evaluate strategies to manage program uncertainty through integration of program metrics and technical measurement with program risk management in accordance with Earned Value Management standards, the DoD Risk Management Guide, and the Defense Acquisition Guidebook.**
	Recognize the role of systems engineering in establishing a cost, schedule, and performance baselines for a given project.
	Describe a process for developing a set of program metrics and measures (EVM and TPMs) to support integrated technical assessment and program management.
	Identify sources of program technical risks.
	Given a specific risk, analyze the risk in accordance with the DoD Risk Management guide.
	Derive a set of technical performance measures to support management of program risk.
	Discuss strategies to integrate use of technical performance measurement and Earned Value Management to manage program risk and uncertainty.
8	**The student will describe major elements of system security engineering as it applies to the design, development, procurement, and sustainment of DoD systems in accordance with the directives and guidance discussed in class.**
	Define system security engineering related terms of critical program information (CPI), mission critical functions, mission critical components, trusted systems and networks (TSN), and supply chain.
	Describe the major elements of program protection plan (PPP).
	Describe the methods for determining system security engineering requirements and design features.
	Describe CPI risk assessment concept.
	Describe horizontal protection process.
	Describe TSN risk assessment concept.
	Recognize the appropriate application of specific countermeasures to the protection of technology, mission-critical functions, mission-critical components, and information.
	Describe countermeasure considerations of foreign involvement to include Anti-Tamper.
	Describe software assurance (SwA).
	Describe supply chain risk management (SCRM).
	Describe cybersecurity as a countermeasure.
	Describe the integrated risk management concept.

Figure 10. ENG 301—Leadership in Engineering Defense Systems Course Objectives

10	**Given a scenario, apply risk management actions and processes for a program.**
	Differentiate between "risks" and "issues."
	Describe the DoD Risk Management Process.
	Identify program risks and their associated root causes.
	Apply risk analysis techniques to a program's risk areas.
	Identify potential risk mitigation options for risk events.
	Describe methods for risk tracking.
	State the objectives of risk planning.
11	**Determine organizational structures to manage risk.**
	Describe typical potential program risk management organizational structures.
	Explain the importance of government and contractor risk management collaboration.
12	**Use risk management software for risk analysis.**
	Determine risk ratings using risk management software.
	Apply Monte Carlo simulation software for a program schedule analysis.
	Examine program risk levels based on the results of Monte Carlo simulations.

Figure 11. PMT 251—Program Management Tools, Part I Course Objectives

1	**Generate program solutions and documents using program management tools and techniques in an acquisition program scenario.**
	Using scenario documentation; develop a Team Charter for an Integrated Product Team (IPT) describing the team's purpose, goals, roles, operating agreements ands critical success factors
	Using a list of product office personnel, construct an IPT addressing both personnel availability and background/experience
	Given market research and company data, assess technology maturity to determine technology readiness levels
	Using product office documentation; select cost, technical, and schedule criteria to evaluate system alternatives
	Using product cost, technical, and schedule data; evaluate alternatives to determine an order of preference
	Given revised scenario data, re-evaluate alternatives to determine a recommended alternative
	Given a simulated contractor's schedule and supporting documentation, analyze them to identify errors
	Using scenario documentation, apply risk management processes to identify and analyze risks, and propose mitigation plans
	Given Earned Value Management (EVM) data, analyze it to determine project status
	Provided additional EVM data, evaluate it to detect trends in contractor performance
	Using EVM data, assess contract performance, to identify significant risks, issues, and recommended courses of action
	Provided simulated contractor cost data and guidance, use cost estimating techniques and factors to develop a detailed cost estimate
	Given scenario information, prepare an information issue paper to document results of issue analysis
	Given a project risk occurrence, apply problem solving tools to determine a recommended course of action
	Using results of issue analysis, develop a decision briefing to justify a recommended course of action
	Given sample risk management software tools, compare their features to determine their advantages and disadvantages
	Given a scenario, apply project management tools to generate project documentation

Figure 12. PMT 257—Program Management Tools, Part II Course Objectives

2	**Given a situation, identify risks in the formulation of a transition or acquisition strategy.**
Explain supply chain management risks to manufacturing.	
14	**Given a scenario analyze a program production plan and its relationship to the overall acquisition strategy.**
	Explain the purpose of manufacturing risk management.
	Define the common manufacturing risk categories.
	Define the 5 M's of Manufacturing.
	Explain the purpose and benefit of the Best Manufacturing Practices Center of Excellence.
	Explain the basics of Lean Manufacturing.
	Explain the risks associated with unstable rates and quantities, variation in processes, and special tools and equipment.
18	**Analyze software measurement data and develop software risk mitigation plans.**
	Given a scenario and the Practical Software Measurement Methodology (PSM), prepare a set of appropriate software management, quality and process measures.
Explain the importance and methods of determining ESOH risk levels.	
Identify issues that impede effective ESOH implementation.	
Explain ESOH risk management techniques.	

Figure 13. PMT 352A—Program Management Office Course, Part A Course Objectives

6	**Given a situation, identify risks in the formulation of a transition or acquisition strategy.**
	Given a contract, prepare an integrated risk assessment in response to a government solicitation.
	Analyze risk items regarding cost, schedule and technical maturity.
	Recognize the technology transition mechanisms and their roles.
	Recognize four practices that support the rapid and effective transition from science and technology to products.
	Given program requirements, the acquisition strategy and risk management software, analyze the program's risks.
21	**Given a scenario with design decisions to document and ensure deliveries reflect a best-value balance among cost, schedule, performance and technical risk factors, relate appropriate system engineering technical management processes that assess alternatives and measure progress.**
	As testing progresses, modify technical risk plans to address unexpected results.
	Conduct trade studies to achieve appropriate balances between cost, schedule, performance, and risk.
	As testing progresses, revise design decisions as required to meet and exceed cost and performance requirements.
	Given a government solicitation, synthesize design alternatives, risk, and requirements.
	Given program status, discuss recommended programmatic changes at an information brief in the context of risk mitigation.
	Assess software quality risks associated with software development for a given program.
	Given a scenario, develop appropriate mitigation strategies for software acquisition risks you identified.
40	**Given a system's status and user feedback, re-evaluate program risk and revise mitigation plans.**
	Propose a program's plan for risk mitigation.
	Given international and interoperability issues, revise risk management plan and risk handling/mitigation plan.
	Given production changes/challenges, assess a program's risk and associated risk handling options.
	Explain the purpose of manufacturing risk management.
	Define the common manufacturing risk categories.
	Explain the risks associated with unstable rates and quantities, variation in processes, and special tools and equipment.

Figure 14. PMT 352B—Program Management Office Course, Part B Course Objectives

10	**Recognize the DoD acquisition risk management process within and IPPD/IPT environment.**
	Recognize current Department of Defense risk management policy for acquisition programs.
	Identify the basic categories and examples of risk for acquisition programs.

Figure 15. PQM 101—Production, Quality, and Manufacturing Fundamental Course Objectives

Describe manufacturing as a design constraint, risk driver, and an enabling/critical set of technologies
Describe producibility as a design constraint, risk driver, and manufacturing enabler
Identify and assess basic supply chain operational issues, problems and risks
Identify and discuss methods, tools and techniques used to manage and mitigate ESOH risks
Identify and describe the key elements of the DoD Risk Process Model
Conduct a Manufacturing Readiness Assessment on an assigned program identifying manufacturing and/or quality risks

Figure 16. PQM 301—Advance Production, Quality, and Manufacturing Course Objectives

Not surprising with little emphasis placed both in policy and education, qualitative data from the Risk and Opportunity Management Survey confirm the perception that DoD acquisition professionals are more familiar with, and that organizations are more actively pursuing, RM over OM.

Quantitative Methodology

A literature search has not revealed any qualitative or quantitative research on the acquisition workforce professionals' perception of their understanding of RM and OM, or their leadership and organization's pursuit of RM and OM.

Research Questions and Hypotheses

The scientific aim of this research is to establish the need to increase the emphasis in education and practice of OM by comparing it to the emphasis on RM. While a comparison between RM and OM provides an evaluative tool for understanding any deficiencies in OM education and/or practice, the author is not advocating any decrease in RM at the expense of OM. Both are valuable tools in program management and cost control. The quantitative research questions directly applicable to the survey are as follows:

1. Do acquisition workforce members perceive that they understand OM equal to RM?
2. Do acquisition workforce members perceive that they work for leadership that encourages OM equal to RM?
3. Do acquisition workforce members perceive that they work for an organization that actively pursues OM equal to RM?

Answering these research questions through use of the survey instrument, focused on RM and OM, the former as a comparative tool, has the potential to provide the DoD with a practical roadmap to increase education in OM and make its practice a useful tool in the pursuit of a should-cost culture.

> Hypothesis 1: Acquisition workforce members perceive that they understand OM less than RM.

The expectation is that acquisition workforce members will perceive that their understanding of OM is less than that of RM. Not only is OM less mature as a discipline, but the current DoD's culture is more focused on ensuring that failure doesn't occur than on creating opportunities for increased success.

> Hypothesis 2: Acquisition workforce members perceive that their leadership encourages OM less than RM.

The expectation is that acquisition workforce members will perceive that, overall, their leadership encourages OM less so than RM. With the RM process being more ingrained in the system engineering process, leaders need to provide greater personal emphasis and expend greater personal energy on OM to create an environment where RM and OM receive equal attention and resources.

> Hypothesis 3: Acquisition workforce members perceive that their organization manages opportunities less than it manages risk.

The expectation is that acquisition workforce members will perceive that their organization manages opportunities less than risk. Again, with RM process being more ingrained in system engineering process, organizations need to have processes that emphasize equal disciple and process control to RM and OM.

Research Design

The most appropriate research design for exploring these hypotheses is to conduct surveys that can capture the perceptions of acquisition workforce members across multiple functional areas, years of experience, services, and organizational type. Questionnaires are an appropriate and relatively easy way to collect information across a wide population for studying behavioral items (Cozby & Bates, 2012). The research design consists of a simple, one-page questionnaire to be distributed via an email with an Opinio-developed questionnaire to previous DAU classes, as well as an identical paper version to be given in multiple DAU residential classes. It is developed specifically to minimize the time required to distribute and complete in order to encourage participation, since the paper version is being conducted using class time. Paper version surveys are post-collection converted electronically into the electronic Opinio format to reap the data collection analysis and reporting tools available through Opinio.

Although preference would be to consider these hypotheses across the total population of interest—that is, all acquisition professionals, including all organizations, services, geographic areas, and experience levels—that would be too costly, impractical, and probably impossible (Acharya et al., 2013). Due to time and cost constraints, and a population of over 100,000 acquisition professionals (45,443 Army, 40,651 Navy, and 25,075 Air Force, when surveyed in 2007), sampling was limited to students attending DAU courses at DAU Midwest campuses, that is, convenience sampling (DAU, 2007). One disadvantage of convenience sampling is that bias can be introduced since a sampling of DAU Midwest students will likely comprise a large percentage of students from the local USAF base, Wright-Patterson AFB, which may not be an accurate representation of the overall acquisition workforce population (Cozby & Bates, 2012).

Population and Sample

Questionnaires were made available to all acquisition workforce students taking classes at DAU Midwest campuses from mid-February through late-March 2017. With an acquisition workforce of over 100,000, a sample size of 384 participants is required to provide a precision of estimates of +5% with a 95% confidence level (Cozby & Bates, 2012). A total of 388 surveys were collected.

Measures/Instrumentation

A survey instrument (see Appendix) was developed specifically for this research. Perceptions on an understanding of, encouragement of, and use of OM was based specific questions on OM, as well as on a comparative analysis of participants' perceptions of their understanding of, leadership encouragement of, and organizational use of RM.

The survey is divided into two sections. The first section includes 10 statements employing a 5-point Likert scale used for each item (1=strongly disagree to 5=strongly agree). The first six questions were directly related to the opportunity to RM comparison. The next three questions were created to qualitatively determine which specific areas of OM were perceptually most encouraged by leadership—continuous improvement in both generic and business processes and reductions in performance requirements. The last question was designed to understand if the acquisition workforce believed that they had specific ideas of where performance requirements could be traded off for financial savings.

In order to better understand the acquisition workforce perceptions on opportunity culture, the following four demographic variable data were collected: branch of service, functional area, years employed in the acquisition workforce, and organizational type.

Data Collection Procedures

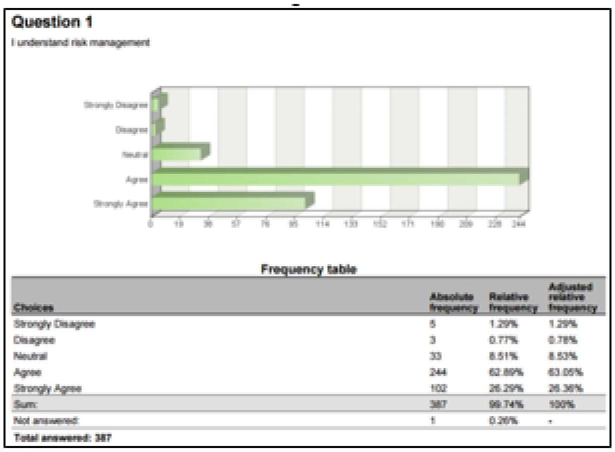
An electronic link was sent to former students directing them to the survey site and ensuring that students understood that the survey was both anonymous and voluntary. Also, with the permission of instructors and consent of the students, the paper questionnaire was given to complete during class, again with the understanding that the survey was both anonymous and voluntary. Students who did not want to participate could either not open the survey, or if given a paper copy, not accept it or return it blank. Upon completion, students anonymously returned the surveys to a table at the front of the room.

Data Analysis

Due to time constraints for both collection and analysis of data, data analysis for each question was limited to between 383 to 388 responses. Also, to ensure that the research was not flawed by an alternative explanation, a succession of t-tests, using a significance level of .05, should be conducted to verify that the data does not vary significantly based on any of the demographic variables. For example, one could theorize that specific functional areas may vary in their perceptions and therefore may bias the data if a significant portion of those respondents were functionally aligned.

The data collected did expose some significant trends in the perceptions of the current acquisition workforce, as the limited number of participants provides for an accuracy of approximately +5% given a 95% level of confidence.

The first pair of questions analyzed were the participants' perceived understandings of both RM and OM. Most notably, analysis of the results from "Question 1: I understand risk management" revealed that 89.2% of the participants either agreed or strongly agreed that they understood RM; whereas, that percentage dropped significantly to 55.9% for those participants that agreed or strongly agreed for the paired question on OM, "Question 4: I understand opportunity management." Figures 17 and 18 provide the details on these two questions.

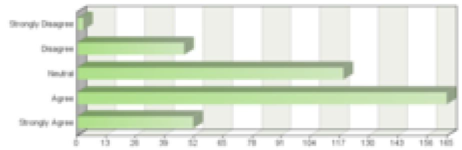


Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	5	1.29%	1.29%
Disagree	3	0.77%	0.78%
Neutral	33	8.51%	8.53%
Agree	244	62.89%	63.05%
Strongly Agree	102	26.29%	26.36%
Sum:	387	99.74%	100%
Not answered:	1	0.26%	-
Total answered: 387			

Figure 17. Question 1 Results

Question 4

I understand opportunity management



Frequency table

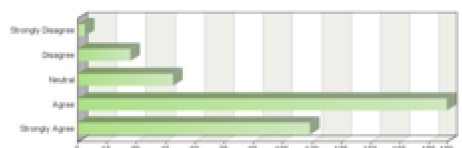
Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	3	0.77%	0.78%
Disagree	48	12.37%	12.4%
Neutral	119	30.67%	30.75%
Agree	165	42.53%	42.64%
Strongly Agree	52	13.4%	13.44%
Sum:	387	99.74%	100%
Not answered:	1	0.26%	-
Total answered: 387			

Figure 18. Question 4 Results

The second pair of analyzed questions were the participants' perception of their leadership's encouragement for them to identify risks and opportunities. Analysis of the results from "Question 2: I am encouraged by my leadership to identify risks" revealed that 79.4% of the participants either agreed or strongly agreed that leadership encouraged them to identify risks; whereas, that percentage dropped significantly to 61.3% for those participants that agreed or strongly agreed for the paired question on OM, "Question 5: I am encouraged by my leadership to identify opportunities." Figures 19 and 20 provide the details on these two questions.

Question 2

I am encouraged by my leadership to identify risks



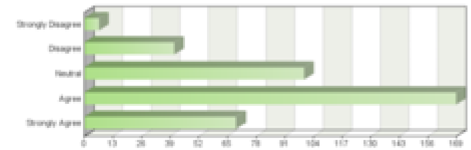
Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	4	1.03%	1.03%
Disagree	27	6.96%	6.96%
Neutral	49	12.63%	12.63%
Agree	189	48.71%	48.71%
Strongly Agree	119	30.67%	30.67%
Sum:	388	100%	100%
Not answered:	0	0%	-
Total answered: 388			

Figure 19. Question 2 Results

Question 5

I am encouraged by my leadership to identify opportunities



Frequency table

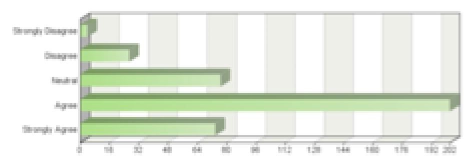
Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	7	1.8%	1.81%
Disagree	41	10.57%	10.62%
Neutral	100	25.77%	25.91%
Agree	169	43.56%	43.78%
Strongly Agree	69	17.78%	17.[illegible]
Sum:	386	99.48%	100.[illegible]
Not answered:	2	0.52%	-
Total answered: 386			

Figure 20. Question 5 Results

The third pair of analyzed questions were the participants' perception of whether or not their organization managed risks and opportunities. Analysis of the results from "Question 3: My organization manages risks" revealed that 71.3% of the participants either agreed or strongly agreed that their organization manages risks; whereas, that percentage dropped significantly to 45.1% for those participants that agreed or strongly agreed for the paired question on OM, "Question 6: My organization manages opportunities." Figures 21 and 22 provide the details on these two questions.

Question 3

My organization manages risks



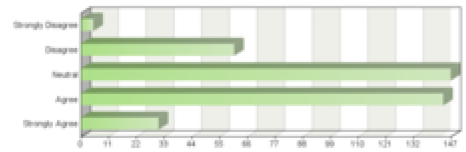
Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	4	1.03%	1.04%
Disagree	27	6.96%	7.03%
Neutral	77	19.85%	20.05%
Agree	202	52.06%	52.6%
Strongly Agree	74	19.07%	19.27%
Sum:	384	98.97%	100%
Not answered:	4	1.03%	-
Total answered: 384			

Figure 21. Question 3 Results

Question 6

My organization manages opportunities



Frequency table

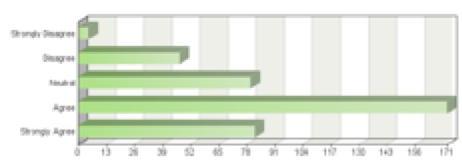
Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	5	1.29%	1.29%
Disagree	61	15.72%	15.72%
Neutral	147	37.89%	37.89%
Agree	144	37.11%	37.11%
Strongly Agree	31	7.99%	7.99%
Sum:	388	100%	100%
Not answered:	0	0%	-
Total answered: 388			

Figure 22. Question 6 Results

The fourth pair of analyzed questions were the participants' perception of whether or not their leadership encouraged process improvement, to include questioning current business processes, in order to save money. Analysis of the results from "Question 7: I am encouraged by my leadership to continuously improve my current processes to save money" revealed that 65.2% of the participants either agreed or strongly agreed that their leadership encouraged their continuous process improvement. "Question 9: I am encouraged by my leadership to question current business process requirements when money can be saved" revealed similar results, with 56.2% of the respondents answering either "agreed" or "strongly agreed." Figures 23 and 24 provide the details on these two questions.

Question 7

I am encouraged by my leadership to continuously improve my current processes to save money



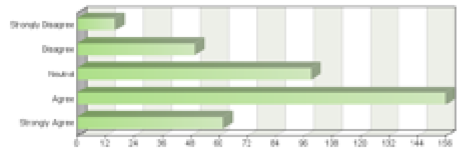
Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	5	1.29%	1.3%
Disagree	47	12.11%	12.21%
Neutral	80	20.62%	20.78%
Agree	171	44.07%	44.42%
Strongly Agree	82	21.13%	21.3%
Sum:	385	99.23%	100%
Not answered:	3	0.77%	-
Total answered: 385			

Figure 23. Question 7 Results

Question 9

I am encouraged by my leadership to question current business process requirements when money can be saved



Frequency table

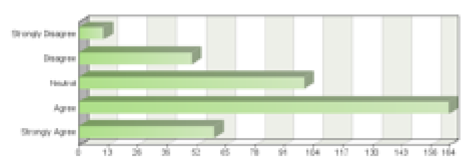
Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	16	4.12%	4.18%
Disagree	50	12.89%	13.05%
Neutral	99	25.52%	25.85%
Agree	156	40.21%	40.73%
Strongly Agree	62	15.98%	16.19%
Sum:	383	98.71%	100%
Not answered:	5	1.29%	-
Total answered: 383			

Figure 24. Question 9 Results

The fifth pair of analyzed questions were the participants' perception of whether or not their leadership encouraged them to question performance (i.e., users) requirements, and secondly, if they in fact could think of "at least one performance requirement" that could be traded in order to save money. Analysis of the results from "Question 8: I am encouraged by my leadership to question performance requirement when money can be saved" revealed similar results with the process improvement paired questions in that 57.7% of the participants either agreed or strongly agreed that their leadership encouraged questioning performance requirements. Question 10 strayed from the previous three questions in that it was ascertaining their opinion versus their perception of leadership. "Question 10: I can think of at least one performance requirement that we could trade-off to save our program money" indicated similar results, with 59.0% of the respondents answering either "agreed" or "strongly agreed." Figures 25 and 26 provide the details on these two questions.

Question 8

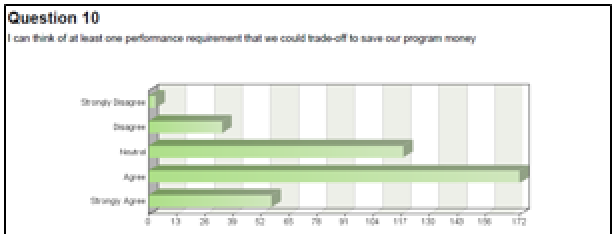
I am encouraged by my leadership to question performance requirements when money can be saved



Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	11	2.84%	2.86%
Disagree	50	12.89%	12.99%
Neutral	100	25.77%	25.97%
Agree	164	42.27%	42.6%
Strongly Agree	60	15.46%	15.58%
Sum:	385	99.23%	100%
Not answered:	3	0.77%	-
Total answered: 385			

Figure 25. Question 8 Results



Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Strongly Disagree	3	0.77%	0.78%
Disagree	34	8.76%	8.85%
Neutral	118	30.41%	30.73%
Agree	172	44.33%	44.79%
Strongly Agree	57	14.69%	14.84%
Sum:	384	98.97%	100%
Not answered:	4	1.03%	-
Total answered: 384			

Figure 26. Question 10 Results

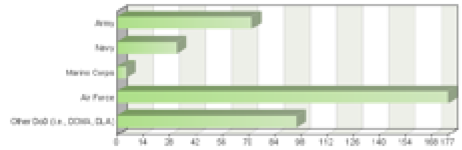
Demographics

As expected due to the survey's administration being primarily performed at the DAU Midwest's Kettering, OH, campus in close proximity to Wright-Patterson AFB, nearly half (45.6%) of the respondents self-identified as "Air Force," compared to only 1.3% self-identifying as "Marine Corps" and 8.3% as "Navy." In using these results for cultural inquiry, one could posit that this skew may affect the overall understanding of the general DoD population as each service may have its own distinct culture. Percent comparisons of individual questions within these groups were not conducted as the amount of data required for +5% accuracy given a 95% level of confidence diminishes the ability to provide definitive quantitative results for each demographic group.

With the exception of the "Contracting/Financial Management" acquisition career field, a fairly even distribution of career fields were represented, with over half being either "Program Management" (25.8%) or "Engineering/Production, Quality and Manufacturing" (25.3%). Likewise, "years working in an acquisition career field" showed a reasonably even distribution. Of the four categories only "2–5 years" at 16% was relatively slightly represented. Finally, the organizational type that the respondents currently work in showed a relatively large percentage engaged in program offices at 40% with a relatively even distribution amongst the remaining four categories; albeit, "research laboratory" was only represented by 5.4% of the participant populations. Figures 27–30 contain detailed results of the four demographic questions

Question 11

Branch of service:



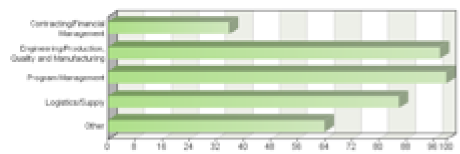
Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Army	72	18.56%	18.85%
Navy	32	8.25%	8.38%
Marine Corps	5	1.29%	1.31%
Air Force	177	45.62%	46.34%
Other DoD (i.e., DCMA, DLA)	96	24.74%	25.13%
Sum:	382	98.45%	100%
Not answered:	6	1.55%	-
Total answered: 382			

Figure 27. Question 11 Results

Question 12

My primary aquisition career field is:



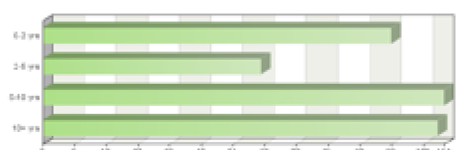
Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
Contracting/Financial Management	36	9.28%	9.38%
Engineering/Production, Quality and Manufacturing	98	25.26%	25.52%
Program Management	100	25.77%	26.04%
Logistics/Supply	86	22.16%	22.4%
Other	64	16.49%	16.67%
Sum:	384	98.97%	100%
Not answered:	4	1.03%	-
Total answered: 384			

Figure 28. Question 12 Results

Question 13

Years working in acquisition career field



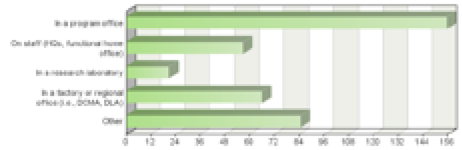
Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
0-2 yrs	99	25.52%	25.58%
2-5 yrs	62	15.98%	16.02%
5-10 yrs	114	29.38%	29.46%
10+ yrs	112	28.87%	28.94%
Sum:	387	99.74%	100%
Not answered:	1	0.26%	-
Total answered: 387			

Figure 29. Question 13 Results

Question 14

I currently work:



Frequency table

Choices	Absolute frequency	Relative frequency	Adjusted relative frequency
In a program office	156	40.21%	40.52%
On staff (HQs, functional home office)	57	14.69%	14.81%
In a research laboratory	21	5.41%	5.45%
In a factory or regional office (i.e., DCMA, DLA)	66	17.01%	17.14%
Other	85	21.91%	22.08%
Sum:	385	99.23%	100%
Not answered:	3	0.77%	-
Total answered: 385			

Figure 30. Question 14 Results

Results

All three hypotheses are supported. Respondents' perceptions, as derived from Question 1 and Question 4, indicate a persistent gap in their own understanding of OM versus RM, which supports the notion and documentation presented earlier that there exists a lack of education on OM, especially when compared with RM, indicated by the wide disparity of 33.3% less respondents specifying that they understood OM (55.9%) compared to RM (89.2%). Also evident from the data (Question 2 and Question 4) is the respondents' perception that their leadership encourages the pursuit of OM less so than RM, by a difference of 18.1% when comparing OM (61.3%) versus RM (79.4%). Finally, data indicates that respondents perceive organizations not managing opportunities as aggressively as risks, denoted by a difference of 26.2% when comparing OM (45.1%) and RM (71.3%). These results supporting the second and third hypotheses corroborate the notion that cultural inculcation of OM has yet to be achieved.

Questions 7–10 reveal some important information regarding the respondents' perceptions on two potential sources of opportunities for cost savings. While the majority of respondents signified that they either agreed or strongly agreed with leadership encouragement or their own ability to identify these opportunities, there is still work to be done if the DoD acquisition workforce, as a unified whole, adopts "should cost" through the grassroots efforts of OM, as large percentages of respondents either were neutral, disagreed, or strongly disagreed with the notion that their leadership encouraged general process improvement (34.0%), business process improvements (42.5%), or questioning performance requirements (41.5%). When questioned about their own ability to identify at least one requirement tradeable for savings, 39.9% selected neutral, disagreed, or strongly disagreed.

The reminder of this paper will offer rationale and ideas in the development of increased education, leadership, and organization engagement and encouragement of OM through continuous process improvement and performance requirement tradeoffs, to foster an OM mindset furthering the potential to create a "should-cost" culture across the DoD.

Increased Education on Opportunity Management to Foster "Should-Cost" Mentality

The large gap between those participants who stated that they understood OM versus those who stated that they understood RM, couple with the demonstrated lack of targeted curriculum on OM, as demonstrated by their absence from DAU's DAWIA courses' learning objectives, makes a strong case for an adjustment. If DoD leadership desires a culture that targets "should cost" versus "will cost," the skills and emphasis should be developed for OM early within the education of our acquisition workforce.

Currently, engineering (ENG), program management (PM), and production, quality, and manufacturing (PQM) functional areas have modules embedded in the residential courses dedicated to RM; however, none have been identified for OM. For example, the advanced PQM course, PQM301, has a module dedicated to Manufacturing Readiness Assessment and Risk Management; however, that module, and nowhere else in the course addresses OM. Both the intermediate (DAWIA Level II) and advanced (DAWIA Level III) courses for PM, PMT257, and PMT360, respectively, have RM as a major curriculum topic, but they make no mention of OM. Yet, DAU cannot unilaterally change the curriculum, because it must create and teach curriculum in agreement with the learning objectives established by the Functional Integrated Product Teams (IPTs) at the OSD. The Functional IPTs will need to direct DAU through the adjustment of learning objectives if OM will be taught to the acquisition workforce in order for the gap in understanding between what RM is versus what OM is to evaporate.

As an interim step, DAU does have the ability to develop a continuous learning objective on OM. Currently, their catalog lists a CLM on RM, but none exists for OM. However, a search on the DAU website for "opportunity management" reveals an entry in *Acquipedia* and a "hot topic" recorded in June 2016 on Risk and Opportunity Management.

Process Improvement

Two survey questions, 7 and 9, were developed specifically to gauge the acquisition workforces' perception of their leadership's encouragement of their pursuit of process improvement, generally and specifically to business processes, which is considered a key for should-cost implementation.

The role of "lean" in developing a "should-cost" culture was identified as implementation strategies by both USD(AT&L) Carter and the Air Force Service Acquisition Executive, David Van Buren. Shortly after the release of BBP 1.0, Carter and Mueller (2011), in their *Defense AT&L Magazine* article, *Should Cost Management: Why? How?,* assert that PMs should "call in the assistance of Lean Six Sigma experts to assess your processes and trim the fat. Encourage your contractors to similarly self-evaluate and jointly look at inefficiencies in processes you engage in together" (p. 18). Van Buren and the Assistant Secretary of the Air Force (Financial Management and Comptroller) Jamie Morin stated in their memorandum, *Implementation of Will-Cost and Should-Cost Management*, that "program managers must begin to drive leanness into their programs by establishing Should-Cost estimates at major milestone decisions" (Morin & Van Buren, 2011). The two survey questions addressing process improvement, as previously reported, indicate that the majority of respondents felt that their leadership encouraged them to both "continuously improve [their] current processes to save money" (65.1%) and to "question current business process requirements when money can be saved" (56.2%). While this is encouraging, a nearer uniform attitude toward these pursuits would be highly beneficial.

While specific implementation of lean is beyond the scope of this research, multiple gurus of lean philosophy have written of the importance of lean being inculcated into the culture rather than a program to follow. Jeffery Liker (2004), author of *The Toyota Way*, has indicated that "most attempts to implement lean have been fairly superficial. The reason is that most companies have focused too heavily on tools such as 5S and just-in-time, without understanding lean as an entire system that must permeate an organization's culture."

James Womack, credited with expanding the awareness of the power of lean via his co-authored book with Daniel Jones and Daniel Roos, *The Machine That Changed the World*, states, “The big danger is that it becomes a ‘program’ that everyone is doing as a staff exercise but which no one understands and no one believes in. Then it is just another collection of tools without a context. It inevitably will fail” (Industry Week, 2005, p. 5). This paper’s author’s personal experience as senior manager of Continuous Improvement and Manufacturing while employed in the defense industry also attests to this, writing in his article, *Lean Implementation: A Three-Pronged Attack*, “it became apparent that if we were to successfully attain an attitude of continuous improvement—faster, better, cheaper, we needed to create a culture that would allow lean to thrive” (Riel, 2012, p. 35).

Using OM as a disciplined tool to encourage critical thinking for program cost reduction presents an increased opportunity to drive culture change, provided it becomes more than a “staff exercise,” and instead is used to drive culture change.

Performance Requirements Tradeoff Opportunities

In 2015, the Air Force announced its “Bending the Cost Curve” initiative, introducing a cost-capability analysis program, which offers cost-savings opportunities. As William LaPlante, former Assistant Secretary of the Air Force for Acquisition (SAF/AQ), explains, “The warfighter can point us to the knee in the curve and say, ‘You know what? I’m not willing to pay more for this capability than that capability’” (Serbu, 2016, p. 2). LaPlante further explains that the Air Force “will match up the costs of each individual capability within a set of requirements and ask end users whether the mission benefit is worth the price” (Serbu, 2016, p. 2). However, the purpose for this discussion is pre-contract award, designed to “determine how the Air Force approaches its source selection” (p. 2). As LaPlante states, “Now we can put together a cost capability RFP that shows exactly where we’re willing to pay extra to get a capability” (Serbu, 2016, p. 2). While this initiative is very worthwhile and can surely allow the defense industry a better understanding of where the warfighter places priorities in dollar terms, it does not extend post-contract, when risks and opportunities become more evident as the detailed design progresses.

The discussion on how much a certain requirement threshold is worth needs to extend beyond contract award. By offering a reliable way for a steady, long-term requirements review by the systems engineering community as the design progresses, the warfighters can be provided with a better understanding of optional short and long-term cost avoidance and/or savings prospects, that is, should cost that may become apparent as the design matures (Riel, 2017). Using OM methodology, a first step would be to outline the cost-avoidance/savings opportunity using robust “tradeoff–benefit” statements, employing practices from the RM section of the 2017 RIO Guide. Next, decide upon, track, and report on these opportunities as any other opportunity created in a more robust OM process than currently apparent in organizations.

Although the RIO Guide presents opportunities as more geared towards investing financially today to improve future benefits, directing incremental requirements compromises today to lessen current risks and achieve future benefits is undoubtedly in the crux of the process, as the RIO Guide clarifies to “not ignore small improvements,” which when combined can prove critical to the cost avoidance and/or savings of the requirements tradeoff process (Riel, 2017). Using the models in the RIO Guide for registry development, requirement tradeoff opportunities (RTOs) can be documented and tracked using the same handling choices outlined in the RIO Guide—pursue now, defer, reevaluate, or reject. Also, employing a parallel procedure as found in the RIO Guide, a Requirement Tradeoff Opportunity Register could be created that designates the RTO; the likelihood of the warfighter community reducing or eliminating the requirement; the negative impact on performance; and any positive impact on producibility, reliability, maintainability, and life-cycle costs (Riel, 2017). This extension of the OM process is only possible if the OM process itself enjoys more consistent application as a tool for should cost.

Using Opportunity Management to Drive Culture Change

Although inculcating a should-cost culture may include practical aspects like better education and increased employee knowledge, skills, and motivation, successful transformation lies in the changes to its culture (Morgan, 2006). Culture change within the DoD can be difficult due to its hierarchical control and mechanistic structure. Yet, these challenges can be overcome with a shared, articulated vision via cultural values and artifacts that encourage the acquisition workforce to convert from a heavy RM bias to a more balanced RM/OM approach. Robust organizational cultures "generate an almost tangible social force field of energy that empowers employees" (Ojo, 2010, p. 4) and are associated with increased performance. While the basic assumptions' layer of culture will not change as progress is sought from inherently risk-based to include as an equal opportunity-based, espoused values articulated through visible artifacts, including opportunity registers, signage, Opportunity Management Reviews, and so forth, will need to change for organization culture change to occur. The second layer of organizational culture, espoused values, provides the mission, goals, standards, and other measures designed to shape the organization's strategies, decision-making, and leadership behaviors (Duke & Edet, 2012). Although DAU education will be important for the workforce to understand what OM is, the switch from predominately risk-based to an equal risk/opportunity-based model, requires leadership to ensure an added emphasis on OM as currently experienced by RM. Organizations tend towards stability as leaders intuitively seek to reduce risk through controls and structure, yet, at the "price of diminished innovation and zeal" (Jain, 2013, p. 106), which will be required for OM and should cost to succeed. The addition of OM as an equal partner to RM and the driver for a "should-cost" mentality demands that "the shift goes all the way to [the] core of the culture" (Kofman & Senge, 1993, p. 17). However, Morgan (2006) elucidates that leaders who "understand the challenge of culture change recognize the enormity of [the] task" (p. 138). Culture is not something easily swayed, but rather needs to be cultivated over time. The importance and employment of OM will need to be championed repetitively through the use of artifacts. Formalization of new values coupled with consistency between words and actions can drive trust and create an atmosphere conducive to change (Michailova, 2000). Artifacts—such as posters, brochures, and published stories of OM successes—can add to the inculcation of OM to the DoD culture.

Regarding the effects of organizational culture on change, Hatch and Cunliffe (2013) cite the research of Dan Denison, who "proposed that an organization's strategy, culture and environment need to be aligned if an organization is to achieve high performance" (p. 186), judging that if culture affects behavior, then by managing the culture, preferred behaviors will develop. Leaders will need to incentivize the transition of the culture from one that emphasizes RM to one that allows system engineering's attention to be shared with OM (Fairbanks, 2006). One key to instilling OM is to ensure that the progress and results from the established initiatives are tracked. Husband (2014) describes this process as potentially the "most important step," stating that "without a tracking mechanism and a means to evaluate results, the efforts to create and develop plans for Should Cost initiatives are likely to be wasted" (p. 578), thus illustrating another reason for establishing such initiatives in an opportunity register. Willen and Garber (2011) advocate the use of a "detailed action plan with metrics that can gauged at specific milestones, starting with an aggressive implementation "mindset" to ensure that the SCR [Should Cost Review] is not viewed as merely a study that when completed ends up on a shelf, unused," and that "an aggressive attitude for challenging the status quo" is in place. That will take leadership.

Leadership and the Creation of a "Should-Cost" Culture

By using Ouchi's concept of clan control, where new members are socialized into the culture and thus internalize the DoD's values, principles, and purpose, the members of DoD leadership who control reward, recognition, and promotions can heavily influence the behavior and direction that the DoD adopts (Hatch & Cunliffe, 2013). However, an overemphasis on competition, where short-term results and looking good can trump performance over the long haul, can produce short-term results detrimental to long-term success of the weapon system (Kofman & Senge, 1993). Even beneficial long-term organizational goals must be coupled with ethical leadership, as creating hard, specific should-cost goals does not come without risks, as negative side effects can emerge. For example, reaching current acquisition phase, should-cost goals should not come at the expense of life cycle costs or required operational needs. Demanding goals can challenge ethical behavior, as narrow focus and ambition may cause program management teams to fixate on accomplishing specific should-cost goals without regard for greater DoD organizational goals and values. This is not to say that pride in accomplishing should-cost and OM goals is all bad; as Locke and Latham (2009) clarify, the possible drawbacks of goal-setting can be alleviated by managerial attention and solid, ethical leadership, citing that "organizations cannot thrive without being focused on their desired end results any more than an individual can thrive without goals to provide a sense of purpose" (p. 22).

Conclusions and Recommendations

As Husband (2014) points out, experienced acquisition officials will note that should cost and OM are not new; however, they will "require an abundance of strategic thinking and planning, and a long-term vision" (p. 589). In other words, developing a should-cost culture and employing some of the tools and opportunities described herein will require a leadership-centered approach. It will also need to engage the hearts and minds of the entire acquisition workforce, down to the grassroots level. Increased education in managing opportunities, to include requirement tradeoff opportunities, can and should play a role in the development of OM at the lowest levels of an organization, so that an overall "should-cost" culture can develop and endure. However, organizational change is difficult. Leaders wanting to produce a culture that gives equal precedence to OM as to RM face a formidable challenge and should recognize the work that it will take (Ivancevich et al., 2011; Morgan, 2006).

More complete quantitative research and longitudinal studies are recommended to understand whether the acquisition workforce trends towards an equitable assumption of OM and RM responsibilities over time as an indicator of culture shift success. Requirements should be challenged during the entire design process, using an RTO mentality, to ensure that they retain their value as the design matures. Leadership must create reward and recognition mechanisms consistent with reaping opportunities versus bias toward risk control to facilitate the shift towards a more balanced RM/OM resource allocation. Cultural markers, such as should-cost and OM success stories, need to become persistent artifacts to help change the culture. The 2016 *Performance of the Defense Acquisition System Annual Report* is correct in its statement that "the institution of 'should cost' management and its consistent emphasis over the last 6 years by the acquisition chain-of-command has been a success and should be a permanent feature of the DoD's acquisition culture" (pp. xviii, xix). However, it will take a leadership-driven, persistent emphasis to change that culture.

References

Acharya, A. S., Prakash, A., Saxena, P., & Nigam, A. (2013). Sampling: Why and how of it? *Indian Journal of Medical Specialties, 4*(2), 330–333

Boito, M., Brancato, K., Graser, J., & Cook, C. (2012). *The Air Force's experience with should-cost reviews and options for enhancing its capability to conduct them*. RAND.

Burt, D. N. (1972, September–October). "Should cost" a multi-million dollar savings. *Air University Review.*

Carter, A. B., & Hale, R. (2011). Joint memorandum on savings related to "should cost." Washington, DC: DoD.

Carter, A. B., & Mueller J. (2011, September–October). Should cost management: Why? How? *Defense AT&L Magazine,* 14–18.

Conrow, E., & Charette, R. (2008, March–April). Opportunity management: Be careful what you ask for. *Defense AT&L Magazine*, 16–19.

Cozby, P. C., & Bates, S. C. (2012). *Methods in behavioral research* (11th ed.). New York, NY: McGraw-Hill.

Defense Acquisition University. (2007). *Defense acquisition structures and capabilities review*. Washington, DC: DoD.

DoD. (2017, January). *Risk, issue, & opportunity management guide for defense acquisition programs.* Washington, DC: Author.

Dester, W. S., & Blockley, D. I. (2003). Managing the uncertainty of unknown risks. *Civil Engineering and Environmental Systems, 20*(2), 83–103.

Duke, J., & Edet, G. H. (2012). Organizational culture as determinant of non-governmental organization performance: Primer evidence from Nigeria. *International Business and Management*, (1), 66.

Fairbanks, W. P. (2006). Implementing transformation vision. *Joint Forces Quarterly, 42*(3), 36–42.

Fiol, C. M., & Lyles, M. A. (1985). Organizational learning. *The Academy of Management Review, 10*(4), 803. Retrieved from http://search.proquest.com/docview/210944840?accountid=40390

Hatch, M., & Cunliffe, A. L. (2013). *Organization theory: Modern, symbolic, and postmodern perspectives/Mary Jo Hatch with Ann L. Cunliffe.* Oxford, UK: Oxford University Press.

Husband D. M. (2014, April). Applications of should cost to achieve cost reductions. *Defense Acquisition Review Journal, 21*(2), 565–594

Industry Week. (2005). Lean manufacturing visionary James Womack on frontiers of lean thinking [Q&A from webcast]. Retrieved from https://www.lean.org/Downloads/IW.pdf

Ivancevich, J. M., Konopaske, R., & Matteson, M. T. (2011). *Organizational behavior and management* (9th ed.). Boston, MA: McGraw-Hill/Irwin.

Jain, P. (2013). Confessions of an Enron executive: We lacked finesse. *Emergence: Complexity and Organization, 15*(2), 104–109.

Kofman, F., & Senge, P. M. (1993). Communities of commitment: The heart of learning organizations. *Organizational Dynamics, 22*(2), 4–23.

Liker, J. (2004). *The Toyota way: 14 management principles from the world's greatest manufacturer.* New York, NY: McGraw-Hill.

Locke, E. A., & Latham, G. P. (2009). Has goal setting gone wild, or have its attackers abandoned good scholarship? *Academy of Management Perspectives, 23*(1), 17–23.

Maucione, S. (2017). DoD report shows big savings from "should cost" policy. *Federal News Radio*. Retrieved http://federalnewsradio.com/defense/2017/01/dod-report-shows-big-savings-cost-policy/

Mechanistic organizations. (n.d.). In *Reference for Business, Encyclopedia of Business* (2nd ed.). Retrieved from http://www.referenceforbusiness.com/management/Mar-No/Mechanistic-Organizations.html

Michailova, S. (2000). Contrasts in culture: Russian and Western perspectives on organizational change. *Academy Of Management Executive, 14*(4), 99–112.

Morgan, G. (2006). *Images of organization/Gareth Morgan*. Thousand Oaks, CA: Sage.

Morin, J., & Van Buren, D. (2011). *Implementation of will-cost and should-cost management* [Memorandum for Record]. Department of the Air Force.

Ojo, O. (2010). Organisational culture and corporate performance: Empirical evidence from Nigeria. *Journal of Business Systems, Governance & Ethics, 5*(2), 1–12.

Pridgen, R. D., Mallon, P., Mallicoat, D., & Triplett, J. (2012, May–June). *Opportunity management: Return on investment—Realized. Defense AT&L Magazine*, 17–20.

Riel, D. (2012, July–August). Lean implementation: Three-pronged attack. *Defense AT&L,* 34–36.

Riel, D. (2017). User requirements: An enduring conversation. *2016 Major General Harold J. "Harry" Greene Awards for Acquisition Writing* (pp. 6–9).

Risor, D. (2011). "Should cost" vs. "did cost": How the military-industrial complex swindles billions of our dollars. *truthout*. Retrieved from http://www.truth-out.org/news/item/4683:should-cost-vs-did-cost-how-the-militaryindustrial-complex-swindles-billions-of-our-dollars

Serbu, J. (2015). Air Force claims $2 billion in acquisition savings from "should-cost" management. *Federal News Radio*. Retrieved from http://federalnewsradio.com/defense/2015/02/air-force-claims-2-billion-in-acquisition-savings-from-should-cost-management/

Under Secretary of Defense for Acquisition, Technology, and Logistics (USD[AT&L]). (2016). *Performance of the Defense Acquisition System, 2016 Annual Report*. Washington, DC: DoD.

Whittinghill, C. (2011). *An evaluation of the perceived organizational culture and innovative climate of a Department of Defense community of organizations* (Order No. 3462071). Available from ABI/INFORM Complete (875885915). Retrieved from http://search.proquest.com/docview/875885915?accountid=40390

Willen, B., & Garber, R. (2011). *Should-cost review: A pragmatic approach to affordability*. A. T. Kearney.

Appendix. Risk and Opportunity Management Survey

Risk and Opportunity Management

Thank you for participating in this short survey on your understanding and experience working with risk and opportunity management. Your candid responses will help us understand the DoD's cultural inculcation of these two important topics.

1. **I understand risk management**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

2. **I am encouraged by my leadership to identify risks**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

3. **My organization manages risks**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

4. **I understand opportunity management**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

5. **I am encouraged by my leadership to identify opportunities**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

6. **My organization manages opportunities**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

7. **I am encouraged by my leadership to continuously improve my current processes to save money**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

8. **I am encouraged by my leadership to question performance requirements when money can be saved**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

9. **I am encouraged by my leadership to question current business process requirements when money can be saved**
   ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

10. **I can think of at least one performance requirement that we could trade-off to save our program money**
    ○ Strongly Disagree ○ Disagree ○ Neutral ○ Agree ○ Strongly Agree

<u>Demographics</u>

11. **Branch of service:**
    ○ Army ○ Navy ○ Marine Corps ○ Air Force ○ Other DoD (i.e., DCMA, DLA)

12. **My primary aquisition career field is:**
    ○ Contracting/Financial Management ○ Engineering/Production, Quality and Manufacturing ○ Program Management ○ Logistics/Supply ○ Other

13. **Years working in acquisition career field**
    ○ 0-2 yrs ○ 2-5 yrs ○ 5-10 yrs ○ 10+ yrs

14. **I currently work:**
    ○ In a program office ○ On staff (HQs, functional home office) ○ In a research laboratory ○ In a factory or regional office (i.e., DCMA, DLA) ○ Other

Finish

THIS PAGE INTENTIONALLY LEFT BLANK

Panel 14. Reducing Life Cycle Costs Through COTS and Additive Manufacturing

Thursday, April 27, 2017	
11:15 a.m. – 12:45 p.m.	**Chair: Captain Eugene Cash, USN**, DDG 1000 Program Office ***Shrinking the "Mountain of Metal": The Potential of Three Advanced Technologies*** David N. Ford, Texas A&M University Thomas J. Housel, Naval Postgraduate School Johnathan C. Mun, Naval Postgraduate School ***Using Additive Manufacturing to Mitigate the Risks of Limited Key Ship Components of the Zumwalt-Class Destroyer*** LT Xiao Y. Wang, USN LCDR James R. Whitworth, USN ***COTS Fresh Look: Use in Major Weapon Systems' Acquisition—Progress, Challenges and Benefits*** Elisabeth Wright, JLo Consulting

Captain Eugene Cash, USN—a former planning officer in DLA Distribution headquarters office of strategic plans, has been selected for the rank of captain.

Cash served as planning officer from June 2010 to July 2013. One of his major initiatives during this time was the development and institutionalization of DLA's Integrated Distribution Strategy for Outside the Continental United States distribution support. Under IDS, he worked to develop solutions and to produce a level of optimization that would improve visibility, reduce transit time, and ultimately result in higher readiness.

Numerous initiatives under the IDS umbrella produced significant benefits. An example of one such success is DLA Distribution Europe's support to Afghanistan customer requirements, which resulted in an improvement in logistics response time from approximately 32 days to 13 days. Additionally, the movement of CENTCOM support from Kuwait to Bahrain saved more than $50 million in infrastructure costs and allowed DLA Distribution to support an increased number of 5th Fleet units and forward stocking requirements.

In late 2012, Cash was requested to serve within DLA Distribution's operations directorate as executive officer. In his new capacity, he led the execution of worldwide distribution support with 26 distribution centers, improving key performance indicators from 91 to 94% green metrics, and assisting in the achievement of an approximately $14.5 million savings through October 2012 via improved container utilization strategies. Future savings are projected at approximately $80 million.

Cash, a native of San Jose, CA, received his commission through the Navy Reserve Officer Training Corps program at the University of Arizona in May 1994 when he received a Bachelor of Science in Accounting. Additionally, he completed training at the Navy Supply Corps School in Athens, GA.

A qualified Supply Surface Warfare Officer, Cash is also a member of the Defense Acquisition Corps. He earned his Acquisition Professional Community Level 2 certification in Program Management and

Information Technology. He has completed the Joint Professional Military Education Level 1 from the Navy War College.

Shrinking the "Mountain of Metal": The Potential of Three Advanced Technologies

David N. Ford—received his BS and MS from Tulane University and his PhD from the Massachusetts Institute of Technology. He is the Urban Beavers Professor of Construction Engineering and Management in the Zachry Department of Civil Engineering, Texas A&M University. He also serves as a research associate professor of acquisition with the Graduate School of Business and Public Policy at the U.S. Naval Postgraduate School (NPS) in Monterey, CA. Prior to a career in academia, Dr. Ford designed and managed the development of constructed facilities in industry and government. [davidford@tamu.edu]

Thomas J. Housel—specializes in valuing intellectual capital, knowledge management, telecommunications, information technology, value-based business process reengineering, and knowledge value measurement in profit and non-profit organizations. He is a tenured full professor for the Information Sciences (Systems) Department at NPS. He has conducted over 80 knowledge value added (KVA) projects within the non-profit, Department of Defense (DoD) sector for the Army, Navy, and Marines. Dr. Housel also completed over 100 KVA projects in the private sector. The results of these projects provided substantial performance improvement strategies and tactics for core processes throughout DoD organizations and private sector companies. [tjhousel@nps.edu]

Johnathan C. Mun—is a research professor at NPS and teaches executive seminars in quantitative risk analysis, decision sciences, real options, simulation, portfolio optimization, and other related concepts. He received his PhD in finance and economics from Lehigh University. He is considered a leading world expert on risk analysis and real options analysis. Dr. Mun has authored 12 books and is the founder and CEO of Real Options Valuation Inc. [jcmun@realoptionsvaluation.com]

Abstract

Military operations create large amounts of damaged equipment, referred to as "mountains of metal." Traditional and current strategies for shrinking the mountain include shipping much equipment to U.S. depots for repair and overhaul. Three advanced technologies, three-dimensional laser scanning, additive manufacturing, and product lifecycle management, can potentially save costs by relocating and accelerating repair operations. Published forecasts of the evolution of these technologies formed the basis for scenarios of their application to shrinking the mountain at U.S. depots, in-theater support facilities, and at forward stations: current use, near-future use, and distant future use. Knowledge Value Added modeling was applied to four technology adoption scenarios (traditional and the three listed) to the Army's up armor HMMWV fleet to estimate returns on investment for each scenario, costs, and potential savings. Cost savings potential of $1.8 billion in the up armor HMMWV fleet and over $21 billion in operations similar in scale to those in Iraq and Afghanistan are estimated. Conclusions include a recommendation to accelerate the adoption and use of these advanced technologies for equipment repair to shrink the mountain of metal.

Introduction

> "To ensure a high-performing and agile supply chain, DoD materiel managers shall leverage modern technologies … to enhance material management processes."
>
> —*DoD Supply Chain Materiel Management Policy*, Sec. 7a, DoD Instruction No. 4140.01, 2014

Military campaigns such as Operation Iraqi Freedom (OIF), Operations Enduring Freedom (OEF), and the war in Afghanistan required vast amounts of equipment and a substantial supply chain to support operations. For example, over 750,000 end items (e.g., boats, aircraft, vehicles, weapons) valued over $36 billion were deployed in Afghanistan in

2007. The army estimates that it has deployed 40% of its equipment to support OIF and OEF, and the Marine Corps estimated deploying 22% of its total fleet assets in Iraq (Solis, 2006). The Marine Corp estimates that 40% of its ground equipment, 50%–55% of its communications equipment, and 20% of its aircraft equipment were supporting operations (Solis, 2006). Much of this equipment is utilized or damaged, requiring repair. This has created an "enormous" (the GAO's term; Solis, 2006) amount of deployed equipment to be diagnosed and then repaired, overhauled, or disposed.

It is the disposal of this materiel that creates an opportunity for better less costly options. This collection of equipment has been referred to as "the Mountain of Metal," referred to hereafter as the Mountain. Using advanced technologies, that is, additive manufacturing, product lifecycle management and three-dimensional laser scanning technology, a large portion of the waste incurred by this Mountain of metal can be eliminated. This study reviews and quantifies the potential benefits of using these three technologies to reduce the costs of a large portion of this Mountain.

The Army and Marine Corps have similar systems for managing equipment in support of operations. (See Solis, 2006, for parallel descriptions of the two systems.) The Army's system is significantly larger in volume and has been reviewed more extensively. The following is based upon the Army system, with relevant notes concerning the Marine Corps. Conclusions are drawn concerning the cost reductions possible with the acquisition and use of the three advanced technologies to both services.

Although major combat operations ceased in Iraq and Afghanistan as of late 2014, the Mountain remains a major DoD challenge. The DoD's reconstitution process, the process whereby materiel from the Mountain can be certified for reuse making it available again for operational use (GAO, 2016). Figure 1 depicts the components of reconstitution. The Army's reset (the Marine Corps uses the term "recovery") processes are a part of reconstitution and can benefit from the adoption of the three advanced technologies investigated here via a larger percent of reuse of the material in the Mountain.

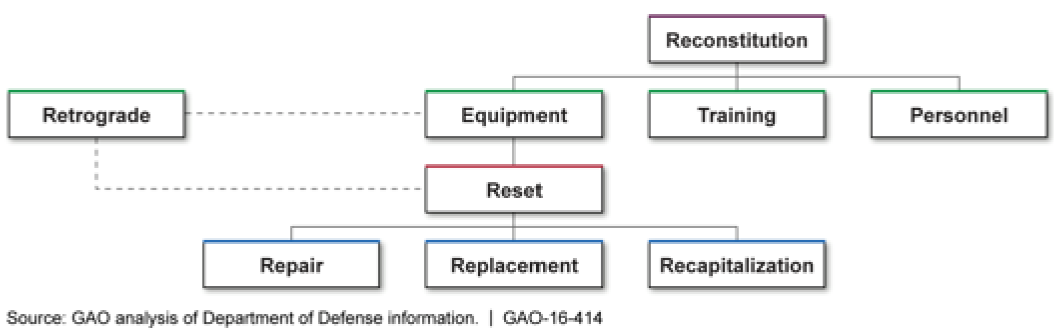


Figure 1. Relationship Between Reconstitution, Retrograde, and Reset Activities
(GAO, 2016)

In theater operations, increased use and harsh operating conditions during operations create the unusable equipment that winds up in the Mountain. Equipment usage rates are several times higher than during peace time[1]. More specifically, the Army reported rates two to eight times higher and the Marine Corps reported rates four to nine times higher than peacetime rates. (Solis, 2006). Gen. Peter Schoomaker, the Army's Chief of Staff, reported to the House Appropriations Subcommittee that "We're wearing out helicopters and trucks, Humvees, tanks at rates that are six, eight, 10 times, in some cases, what we're programmed for." (Hendren, 2007). These usage rates lead to dramatic increases in the costs, not to mention the lack of availability of the equipment, in theater operations.

Making more of the equipment in the Mountain available for reuse would dramatically reduce costs. The Army needs about $13 billion per year for each year of the conflict and for several years thereafter to address the costs of eliminating the Mountain (Hendren, 2007). The Marine Corp costs to eliminate the mountain approaches $1 billion (CBO, 2007).

Processes for Shrinking the Mountain

The *DoD Supply Chain Materiel Management Policy* (2011) specifies five processes by which equipment should be disposed of including how the Mountain can be reduced. In order of decreasing priority, the processes for disposing of materiel from contingent operations are:

1. Consume in theater
2. Reutilize within DoD and other U.S. entities
3. Retrograde (return to U.S. depots) to reset (restore to full capability) U.S. forces
4. Transfer or donate to allies or partner nations
5. Turn-in to DLA Disposition Services for disposal because damage makes reset inappropriate

The efficient repair and overhaul of equipment, using the three advanced technologies, can redirect much equipment for future usage that might otherwise be scrapped (the lowest priority process).

The Traditional Strategy

The traditional Army approach to managing equipment requiring significant maintenance, repair, or overhaul (MRO) is that equipment stays with the unit that it is deployed with and returns to the United States after deployment, where MRO are performed at one of five depots (Figure 2). Some equipment is repaired near forward stations by maintenance companies, reducing transportation costs, saving time, and maximizing availability (FM63-1). However, according to the CBO, "In general, until 2007, Army units rotated in and out of the theater roughly annually, and as a result, most equipment remained in the theater for about a year and was then returned to its unit's home station to be reset [be returned to full capability]" (CBO, 2007). The unit deployed to replace the returning unit

[1] See the Congressional Budget Office (CBO, 2007) study for usage rate details for several types of large equipment and an argument that envisioned Cold War operating tempos should be the benchmark for current operating rates, not peace time tempos.

brought their own equipment. This process was used for hundreds of thousands of pieces of equipment deployed to Iraq, Afghanistan, and surrounding areas (CBO, 2007).

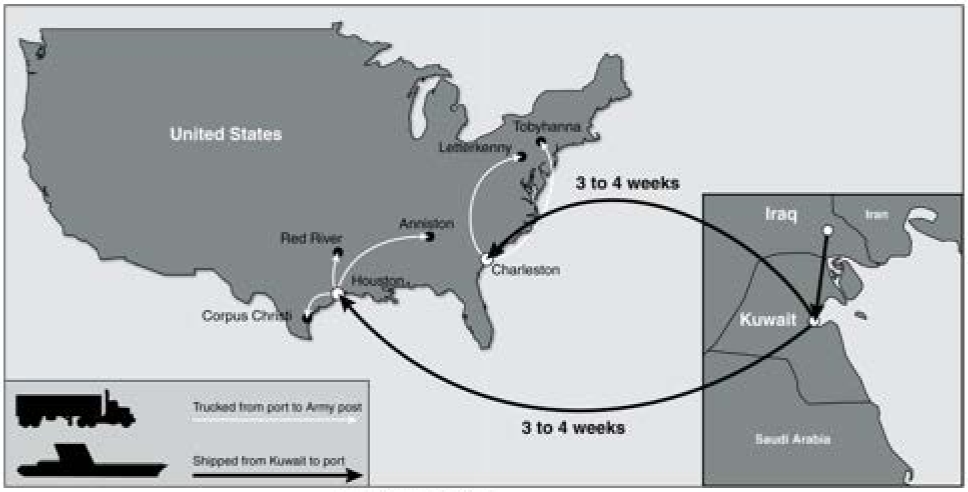


Figure 2. Retrograde of Equipment Leaving Southwest Asia and Returning to the United States for Reset

(GAO, 2012)

The Army uses the reset process to manage damaged equipment. “Reset” is the term for “a series of repair, recapitalization, and replacement actions to restore unit’s equipment to a desired level of combat capability” (Figure 3). This process repairs all damage and performs all routine maintenance (GAO, 2006). Equipment is returned to conditions known as *10/20*, referring to the levels specified by the *10/20* technical manuals which call for all shortcomings and deficiencies to be repaired, and all routine maintenance performed (Taktikz, 2017). Equipment, to be repaired, is often relocated away from forward locations to a reset location through a process referred to as “retrograde” (Aquipedia, 2017). The Marine Corps published a reset implementation plan and the Army published information on aspects of the reset process in 2016 (GAO, 2016).

In-Theater Maintenance, Repair, and Overhaul: The Theater Sustainment Stocks (TSS) and the Theater Provided Equipment (TPE) Initiatives

One disadvantage of the traditional process is that performing repairs in the United States requires transporting the equipment round trip to and from the United States. However, this equipment could be repaired in-theater using the three advanced technologies. The Army initiated two equipment reuse efforts, the Theater Sustainment Stocks (TSS) and Theater Provided Equipment (TPE), in an attempt to increase operational availability and reduce costs. The Theater Sustainment Stocks (TSS) retain an inventory of over 400 types of vehicles and other equipment in theater for deployment with arriving units. The Marine Corp has a similar program named Forward In-Stores. In at least the Army case, this portion of the Mountain typically requires repairs to be operational, and those repairs often do not return the equipment to full capability. For example, the GAO found that less than 7% of a cross-section of ground vehicles in TSS were fully mission capable (Soltis, 2006). Increased in-theater repair capability can increase the operational availability of TSS equipment.

Since its initiation in 2003 the Theater Provided Equipment[2] (TPE) initiative takes force-protection equipment from forces returning to the United States while the equipment is still in theater instead of shipping it back with the units that brought it into the theater. The program transfers the equipment to incoming units. Transfers typically happen at forward stations and departing units are expected to maintain equipment to full mission capabilities. Almost 75% of the Army's trucks in Iraq are in the TPE pool (CBO, 2007). While increasing operational availability of equipment to users and saving shipping costs, the TSS and TPE programs, as currently implemented, prevent depot level MRO such as overhauls. This can require more and more expensive repairs later. Improved MRO in-theater or repairs at forward stations can increase the effectiveness of TPE.

Three Advanced Technologies

Three advanced technologies—that is, three-dimensional laser scanning technologies (3DST), additive manufacturing (AM), and product lifecycle management (PLM)—have the potential to significantly improve the processes used to shrink the Mountain. The following sections provide an overview of these technologies based on a prior study by Housel, Hom, Ford, and Mun, (2015).

Three-Dimensional Laser Scanning Technologies (3DST)

Three-dimensional laser scanning technologies have been used to achieve significant cost savings, optimize maintenance schedules, increase quality, improve safety, and reduce re-work. Commercial applications range from maritime and space applications to manufacturing and production. According to industry analysts, the industry's growth is fueled by the growing recognition that 3D aids in the design, fabrication, construction, operations, and maintenance processes.

Laser scanners use infrared laser technology to produce exceedingly detailed three-dimensional images of complex environments and geometries in only a few minutes. Millions of discrete measurements are captured in every scan. The resulting images, a "point cloud," are millions of 3D measurement points. A complete project may contain hundreds of millions or even billions of points, recreating the complex spatial relationships of the 3D environment. Three-dimensional scanners can be used to get complete or partial 3D measurements of any physical object without any contact with the physical object.

Often used by offshore oil and gas companies to construct and repair oil rigs, 3DST is very effective at documenting oil platforms and refineries to assist in engineering, maintenance, and planning processes. The aerospace and automotive industries have used 3DST for retrofitting floors and measure parts for accurate fit. Other industries using the technology include:

- ***Law Enforcement***. Used in crime scene documentation, forensics and accident reconstruction.

---

[2] Theater Provided Equipment was referred to as "stay behind equipment" until 2005.

- ***Architectural & Civil Engineering***. Used to capture as-built documentation of existing buildings and structures such as bridges provides architects and contractors with exact dimensions. Building Information Models (BIM) can be developed to retrofit projects.
- ***Asset & Facility Management/Documentation***. Three-dimensional documentation of complex factory and plant installations provide users with very precise 3D CAD data for use in facility management, maintenance and asset documentation.
- ***Surveying***. Used to complement or replace traditional tools such as total stations to fully capture manmade or natural objects for volume calculations, as-built surveys and topographic surveys (Faro, 2014).

Additive Manufacturing (AM) (Based on Housel et al., 2015)

Lu, Li, and Tian (2015) contrast AM with equivalent and subtractive forms of manufacturing. Equivalent manufacturing uses the same amount of material to create the product as is in the final product. The mass change during equivalent manufacturing is zero. Casting, forging, and soldering are examples of equivalent manufacturing. Subtractive manufacturing removes material during manufacturing. The mass change during subtractive manufacturing is negative. Milling, turning, and grinding are examples of subtractive manufacturing. In contrast, AM adds material during manufacturing. The mass change in additive manufacturing is positive. Stereolithography is an example of additive manufacturing.

The American National Standards Institute defines additive manufacturing as the "process of joining materials to make objects from 3D model data, usually layer upon layer, as opposed to subtractive manufacturing methodologies" (ASTM, 2013). Additive manufacturing is also commonly referred to as 3D printing. AM differs radically from the currently dominant manufacturing methodologies. Most current methods use subtractive processes (e.g., machining), but AM builds a 3D object by gradually adding successive layers of material that are laid down exactly in their final location. AM does this by fabricating objects directly from 3D computer-aided design (3D CAD) models. The 3D model is disaggregated into multiple horizontal layers, each of which is produced by the machine and added to the preceding layers. Additive manufacturing is often referred to as 3D printing.

AM involves a number of steps from a 3D CAD model to a physical object, as follows:

- **CAD**: A 3D CAD model of the target object is built in software, some times based on a 3D scanned image of the target generated with 3DST. The 3D CAD model determines only the geometry of the target object. The model can be created using 3D laser scanning.
- **Conversion to files for manufacturing**: The CAD model cannot be used directly by AM machines; it must be converted to a format usable by the specific AM technology (e.g., stereolithography) being used. These files describe the external closed surfaces of the original CAD model and forms a basis for calculation of the layers used in manufacturing. The model approximates surfaces of the model with a series of triangular facets.
- **Revision of manufacturing files**: The manufacturing files must be manipulated before manufacturing. For example, multiple objects may be manufactured simultaneously from the same file, requiring that the files of the objects be integrated.

- **Machine setup**: AM machines must be set up to accommodate specific materials, layer thicknesses, and timing.
- **Build**: Although all AM machines follow the layer-by-layer fabrication process, they utilize different techniques and technologies. For example, some of them use a high-power laser beam to melt a very fine metal powder in order to form a thin layer, while some others use UV light to solidify a specific kind of liquid polymer, called *photopolymer*.
- **Post-process**: Post-processing may be required due to the need to cure photopolymers.

The first additive manufacturing system was created in the early 1980s when Charles Hull invented stereolithography (SLA), a printing process that enables a tangible 3D object to be created from digital data. The technology was then used to create a 3D model from a picture and allows users to test a design before investing in a larger manufacturing program. Since then, AM has evolved to include at least 13 different sub-technologies grouped into seven distinct process types.

AM is already a staple in many manufacturing processes and is being increasingly used across a number of industries, including aviation, automobile, and healthcare. Lockheed Martin estimates that some complex satellite components can be produced 48% cheaper and 43% faster with 3D. Production costs could be reduced by as much as 80%. Boeing has installed environmental control system ducting made by AM for its commercial and military aircraft for many years; tens of thousands of AM parts are flying on 16 different production aircraft (commercial and military; Wohlers, 2014). GE Aviation will be using AM will be used to manufacture more than 30,000 fuel nozzles annually for its new LEAP engine starting in 2015. Consolidating 18 parts into one, the new design is 25% lighter and five times more durable than the previous fuel nozzle.

In the automotive industry, Ford Motor Company uses 3D printing in several areas, including the tooling used to create production parts and to build intake manifold prototypes that can be tested for up to 100,000-mile cycles. With traditional manufacturing methods, it would take four months and cost $500,000 to build while a 3D-printed manifold prototype costs $3,000 to build over four days.

Product Lifecycle Management (PLM) (Based on Housel et al., 2015)

The meaning of Product Lifecycle Management (PLM) continues to evolve. It has been defined as an “integrated, information-driven approach comprised of people, processes/practices, and technology, to all aspects of a product’s life, from its design through manufacture, deployment and maintenance—culminating in the product’s removal from service and final disposal. By trading product information for wasted time, energy, and material across the entire organization and into the supply chain, PLM drives the next generation of lean thinking” (Greives, 2006). In another definition by CIMdata, “PLM is a strategic business approach that applies a consistent set of business solutions in support of the collaborative creation, management, dissemination, and use of product definition information across the extended enterprise, and spanning from product concept to end of life-integrating people, processes, business systems, and information. PLM forms the product information backbone for a company and its extended enterprise.” Finally, the Gartner Group defines “PLM is a discipline for guiding products and product portfolios from ideas through retirement to create the most value for businesses, their partners, and their customers.” Although definitions differ, there is agreement that PLM is a systematic approach to managing the series of changes from its design and development to its ultimate retirement or disposal.

A wide range of industries using PLM are finding that 3DLS is becoming a critical tool to link the gap between physical objects in the real world and in the digital design world. The aerospace, automotive, consumer products, manufacturing, and heavy industries all have benefited from faster time to market, improved quality, and reduced warehousing costs with 3D scanning.

Potential Process Options to Shrink the Mountain

Current Capabilities and Forecasted Developments of 3D Scanning Technologies, Additive Manufacturing, and Product Lifecycle Management

A general review of the current and future capabilities of each technology will provide the basis for forecasting how they might be used to shrink the Mountain. The following review of how they might be used immediately and in the future as they add new functionalities is necessary to make reasonable forecasts about how much cost they can reduce over time.

3D Scanning Technologies

Current capabilities and uses of 3D scanning technology include:

- Tabletop scanning and mapping of fixed objects
- Portable, handheld (no mechanical fix to the scanned object) mapping of freeform surfaces (Allard et al., 2013)
- Translation from point cloud collected by scanning to CAD files for design and manufacturing

Potential future capabilities of 3D scanning technology include:

- Scanning technologies integrated with other sensing technologies
- Smart scanning software that automatically diagnoses damage based on scanned data
- Scanners communicating directly with repair facilities
- Scanners communicating directly with manufacturing equipment for automated manufacturing of parts based upon damage assessment
- User-based damage assessment such as units carrying portable 3D scanners for equipment diagnostics

Future applications of 3DST within the DoD can include the use of portable (tabletop-sized) and very portable (handheld) scanners by in-theater repair facilities and at forward stations by repair personnel and equipment users for on-site damage and in-theater assessment and diagnosis. Damage assessment software may be developed to analyze scanned data (e.g., whether actual deviation from design shapes prevents full capability) and thereby speed diagnosis. Three-dimensional scanning technology can be integrated with AM and automated to speed the creation of replacement parts. The technology may eventually be used to sense component conditions while in use and collect user experience data for use in real time conditions assessment and repair.

Additive Manufacturing

Current capabilities and uses of additive manufacturing include:

- Translation from CAD drawings to manufacturing files for use by AM machines
- Making molds for casting parts (Lu et al., 2015)

- Manufacturing with most materials (Lu et al., 2015)
- Manufacturing complex shaped parts (Lu et al., 2015)
- Manufacturing small numbers of parts more cheaply that traditional manufacturing methods (Thomas & Gilbert, 2014)
- Reduction in size of equipment required compared to many traditional manufacturing methods (Lu et al., 2015), allowing more localized manufacturing

Potential future capabilities of additive manufacturing include:

- Redesign the shapes of parts to exploit additive manufacturing advantages for parts such as heat exchangers and lightweight structures (e.g., drone parts; Lu et al., 2015) and custom fitting protective gear (Earls & Baya, 2014)
- Goal-driven computer design of parts that optimizes designs for weight, strength, etc. (Smith, 2015)
- Integrate additive manufacturing into design of part characteristics (Lu et al., 2015). AM can be used to control the internal stresses within a part. Therefore, single parts will, for example, be designed to be stronger at the locations of larger loads.
- Integral design and manufacturing of multiple-material parts (Lu et al., 2015; Smith, 2015). For example, alternating layers of interacting materials with different characteristics such as stiffness and density (Earls & Baya, 2014).
- Manufacturing at the micro and nano scales of objects such as miniature transduces (Lu et al., 2015; Smith, 2015).
- Combination and integration of AM, equivalent, and subtractive manufacturing methods for the manufacturing of parts such as prototypes, molds, electrodes, and casting patterns (Lu et al., 2015)
- Design and use of high-performance alloys such as for high-temperature conditions (Lu et al., 2015)
- Intelligent manufacturing equipment which senses and responds to manufacturing conditions in real time (Lu et al., 2015)
- Consolidation of many components such as sensors, batteries, and electronics into fewer, more complex components, subsystems, and systems. For example, printing circuits, antennas, and RFID tags into products (Earls & Baya, 2014) such as helmets, boots, and clothing (Anusci, 2015).
- Manufacturing of complete subsystems such as small drone wings (Earls & Baya, 2014)
- Small scale and portable manufacturing that allows on-site parts and equipment manufacturing (Smith, 2015)
- Four-dimensional printing in which products change over time in response to conditions, such as for self-assembly, increased strength when in the presence of moisture or a specified temperature (Smith, 2015)

Future applications of AM technologies within the DoD can include their widespread use for making single or small batches of replacement parts from basic materials, manufacturing near forward stations, integration and automation with 3DST for faster parts creation and custom parts, and component designs and manufacturing using diverse and multiple materials, integrated component manufacturing for faster and cheaper repair work, and 4D component design and manufacturing that changes with time or environmental conditions.

Product Lifecycle Management

Current capabilities of product lifecycle management include:

- Aggregation and storage of component-specific data
- Data sharing across user locations and time
- Component life tracking
- Inventory analytics

Potential future capabilities of product lifecycle management include:

- Smart objects that send and receive data and instructions through the PLM system (Shilovitsky, 2016)
- Coordination and communication among connected devices that allow manager-to-component, user-to-component, and component-to-component communication (Shilovitsky, 2016)
- Automated product performance monitoring and reporting in real time (Shilovitsky, 2016)
- User experience data collection in real time and analysis for improved component design (Shilovitsky, 2016)
- Smarter software that can improve repair forecasting and planning by predicting demand (Shilovitsky, 2016)

Future applications of PLM within the DoD can include automated inventory management; repair demand forecasting and planning based on parts conditions; integration of manufacturing across subtractive, equivalent, and additive processes; 4D component design and manufacturing that changes with time or environmental conditions; and the full integration of 3DST, manufacturing, and PLM.

Forecasted Evolutions of the Three Advanced Technologies for Shrinking the Mountain

Advanced technologies uses for shrinking the Mountain are expected to differ by location, that is, whether used at forward stations, in-theater repair facilities, or at U.S. depots. Forecasted applications of each technology in these three locations were developed for three temporal scenarios: current use (Table 1), use in the near future (5–10 years) (Table 2), and use in the distant future (more than 10 years) (Table 3). Location vs. Technology tables with cells describe activities (e.g., maintenance, minor repair, overhaul, and diagnosis

Table 1. Current Repair Applications of Three Advanced Technologies

Location	Current Applications	3D Scanning Technology (3DST)	Additive Manufacturing (AM)	Product Lifecycle Management (PLM)
		Innovative Technology		
Location	US depot	-Limited use for basic parts	-Limited use for basic parts with few materials -Test broader application of basic AM	-Parts and component data storage & sharing -Component life tracking -Inventory analysis
Location	In-Theater	-None or experimental	-Limited use with basic materials	-Limited use
	Forward station	-None	-None	-None

Table 2. Near-Future Repair Applications of Three Technologies

Location	Near-Future Applications	3D Scanning Technology (3DST)	Additive Manufacturing (AM)	Product Lifecycle Management (PLM)
		Innovative Technology		
Location	US depot	-3DST for AM of basic parts replacement is SOP -Test integrated & automated 3DST & AM	-AM for basic parts with basic materials is SOP -Test AM with multi-materials -Test integrated & automated 3DST & AM -Test AM of integrated components	-Automated inventory management is SOP -Test conditional MRO management -Test communication & integration across processes -Test automation across processes -Test providing MRO knowledge & skills with parts
Location	In-Theater	-Damage assessment at micro & nano scales is SOP -Test portable 3DST applications -Test integrated & automated 3DST & AM	-Portable AM is SOP -Test very portable AM -Test integrated & automated 3DST & AM -Test integrated AM & traditional processes	-Test integrated diagnosis & MRO -Test MRO forecasting & planning based on component conditions
	Forward station	-Test real time damage assessment applications -Test very portable scanning applications	-None	-Test real time user experience data collection and use in MRO

Table 3. Distant-Future Repair Applications of Three Technologies

Location	Distant-Future Applications	Innovative Technology		
		3D Scanning Technology (3DST)	Additive Manufacturing (AM)	Product Lifecycle Management (PLM)
Location	US depot	-Scanning for AM for basic parts replacement is SOP -Integrated scanning & AM manufacturing is SOP -Test fully integrated 3DST, AM, PLM	-AM for diverse parts with multi-materials is SOP -Integrated 3DST & AM is SOP -Integrated component AM is SOP -Integrated AM & traditional processes is SOP -Test 4D component design and AM -Test AM-based parts design & micro/nano AM	-Conditional MRO is SOP -Providing MRO knowledge & skills with parts is SOP -Test fully integrated and automated 3DST, AM, PLM
Location	In-Theater	-Integrated scanning & manufacturing is SOP -Portable scanning is SOP	-AM for diverse parts with multi-materials is SOP -Integrated & automated portable 3DST & AM is SOP	-Integrated and automated diagnosis & MRO is SOP -Providing MRO knowledge & skills with parts is SOP
Location	Forward station	-Very portable user scanning damage assessment is SOP -Real time damage assessment & communication to MRO	-Integrated & automated very portable 3DST & AM is SOP	-Integrated diagnosis & MRO is SOP -Providing MRO knowledge & skills with parts is SOP -Real time user experience collection and use in MRO is SOP

Modeling Improved Processes to Shrink the Mountain

We use the knowledge value added methodology to structure the problem of forecasting the future value and cost reductions possible when the three technologies are in place to support shrinking the Mountain. In what follows, we will review the methodology and how it works.

Knowledge Value Added Modeling (Based on Ford et al., 2016)

In the U.S. military context, the Knowledge Value Added (KVA) methodology is a new way of approaching the problems of estimating the productivity (in terms of ROI) for military capabilities embedded in processes that are impacted by technology. KVA addresses the requirements of the many DoD policies and directives by providing a means to generate comparable value or benefit estimates for various processes and the technologies and people that execute them. It does this by providing a common and relatively objective means to estimate the value of new technologies. KVA is a methodology that describes all organizational outputs in common units. This provides a means to compare the outputs of all assets (human, machine, information technology) regardless of the aggregated outputs produced. **It monetizes the outputs of all assets, including intangible knowledge assets.** Thus, the KVA approach can provide insights about the productivity level of processes, people, and systems in terms of a ratio of common units of output (CUO). CUO produced by each asset (a measure of benefits) is divided by the cost to produce the output. By capturing the value of knowledge embedded in an organization’s core processes, employees and technology, KVA identifies the actual cost and value of people, systems, or processes. Because KVA identifies every process required to produce an output and the historical costs of those processes, unit costs and unit values of outputs, processes, functions or services are calculated. An output is defined as the end-result of an organization’s operations; it can be a product or service.

For the purpose of this study KVA was used to measure the value added by the human capital assets and the system assets by analyzing the processes performances. By capturing the value of knowledge embedded in systems and used by operators of the processes, KVA identified the productivity of the system-process alternatives. Because KVA identifies every process output required to produce the final aggregated output, the common unit costs and the common unit values were estimated. KVA quantifies value in two key productivity metrics: return on knowledge (ROK) and return on investment (ROI).

Describing processes in common units also permits, but does not require, market comparable data to be generated, particularly important for non-profits like the U.S. Military. Using a market comparables approach, data from the commercial sector can be used to estimate price per common unit, allowing for revenue estimates of process outputs for non-profits. This also provides a common units basis to define benefit streams regardless of the process analyzed.

Scenarios for Knowledge Value Added Modeling

The three advanced technologies investigated can help shrink the Mountain in three locations: forward stations, in-theater repair facilities, and U.S. depots and at the interactions and integration of repair work at those locations. Figure 3, Processes for Shrinking the Mountain, illustrates the repair process pathways modeled. In what follows, four scenarios were developed that demonstrated the potential cost/benefits of using the three technologies to shrink the Mountain at these three locations.

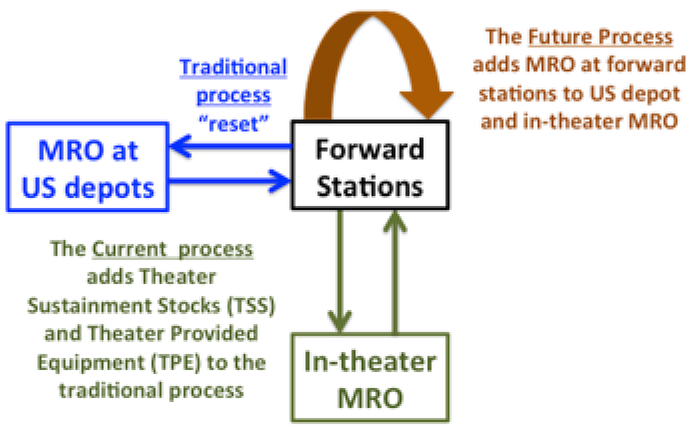


Figure 3. Processes for Shrinking the Mountain

Four advanced technology adoption and use scenarios were developed based on these pathways for modeling the abilities of the three technologies to improve the shrinking of the mountain:

- **The As-Was Scenario** reflects the traditional repair processes, in which all equipment is retrograded from forward stations to U.S. depots, where it is diagnosed, repaired, and overhauled. The equipment is returned to forward stations.
- **The As-Is Scenario** reflects the current processes, which uses the traditional process for some equipment but created Theater Sustainment Stocks (TSS) to provide Theater Provided Equipment (TPE) and in-theater MRO and apply the near-future evolution of the three advanced technologies.
- **The To-Be Scenario** reflects near-future (5–10 years) processes, which will use the traditional processes for some equipment, Theater Sustainment Stocks (TSS) to provide Theater Provided Equipment (TPE) and in-theater

repairs, and forward station repairs for some equipment, using a near-future evolution of the three advanced technologies.

- **The Radical To-Be Scenario** reflects distant-future (more than 10 years) processes, in which all vehicles are diagnosed twice per year, mostly at forward stations and no diagnosis is done at U.S. depots. Simple repairs are performed at forward stations and complex repairs are performed at in-theater facilities. Overhauls are performed at both in-theater facilities and at U.S. depots.

The models were built using the up-armored HMMWV as an example from which extrapolations can be derived to represent the percentage cost/benefits of shrinking the Mountain. This vehicle was chosen because of the relatively large quantity (23,800), their high use in operations (essentially 100% of fleet in Iraq and Afghanistan), and the availability of data. Six variables were used to describe the differences among the four scenarios in the quantitative KVA model, as follows:

- **The number of vehicles that the process was performed on each year at what locations (forward station, in-theater facility, U.S. depot).** For each of the scenarios estimates were made of the fractions of vehicles requiring repair, requiring overhaul, and the fractions of those repairs and overhauls performed at forward station, in-theater, and at U.S. depots. In general, work moved from U.S. depots into in-theater facilities and some then to forward stations over time.
- **Number of times process performed each year per vehicle**. The process frequency for diagnosis and repair at forward stations begins at zero and increases as technology provides means for performing these processes in increasingly difficult circumstances.
- **The average number of employees that performed the process.** In general, the average number of employees required to perform a task decreased with the application of advanced technologies.
- **The average time required to complete the process on a single vehicle**. The average time required to complete a process decreased with the application of advanced technologies.
- **The fraction of the process that is performed using the advanced technologies.** This fraction increased from the traditional to the current and to the near-future scenarios and was largest for the effected processes in the distant future scenario.
- **The cost of the advanced technologies.** The cost of the advanced technologies is partially based on the fraction of automation based on the assumption that partial automation would occur with technology uses as some locations but not others, allowing costs to be controlled.

Results

Returns on Knowledge and Returns on Investment

Table 4 shows the simulated returns on knowledge (ROK) and returns on investment (ROI) of the four scenarios described above.

Table 4. Returns of Simulated Scenarios of Repair of Army's HMMWV Fleet

		--------------Scenario--------------							
		Traditional: As-Was		Current: As-Is		Near Future: To-Be		Future: Radical To-Be	
No.	**Process**	**ROK**	**ROI**	**ROK**	**ROI**	**ROK**	**ROI**	**ROK**	**ROI**
1	Diagnosis at forward station	NA	NA	NA	NA	1612%	1512%	1180%	1080%
2	Repair at forward station	NA	NA	NA	NA	708%	608%	386%	286%
3	Retrograde forward station to US depot	21%	-79%	21%	-79%	15%	-85%	3%	-97%
4	Retrograde forward station to in-theater facility	NA	NA	508%	408%	364%	264%	68%	-32%
5	Diagnosis at US depot	190%	90%	195%	95%	526%	426%	NA	NA
6	Repair at US depot	211%	111%	217%	117%	433%	333%	NA	NA
7	Overhaul at US depot	51%	-49%	52%	-48%	59%	-41%	28%	-72%
8	Transport from US depot to forward station	21%	-79%	21%	-79%	15%	-85%	3%	-97%
9	Diagnosis at in-theater facility	NA	NA	195%	95%	422%	322%	1501%	1401%
10	Repair at in-theater facility	NA	NA	246%	146%	530%	430%	220%	120%
11	Overhaul at in-theater facility	NA	NA	52%	-48%	58%	-42%	28%	-72%
12	Transport in-theater facility to forward station	NA	NA	508%	408%	368%	268%	69%	-31%
	TOTAL	74%	**-26%**	96%	**-4%**	148%	**48%**	169%	**69%**

Note. NA=*Not applicable* because the process is not used in the scenario.

Table 4 also identifies processes that benefit more-or-less relative to each other. The table shows that the diagnosis process, whether performed at forward stations (process #1) or in-theater (process #9), benefits the most from the adoption and use of the three advanced technologies. The ROI for diagnosis increases from 90% in the As-Was scenario and 95% in the current As-Is scenario to over 1400% when performed in-theater in the Radical To-Be scenario.

Table 5 shows the ROK and ROI improvement of the As-Is, To-Be, and Radical To-Be scenarios over the As-Was scenario and the ROK and ROI improvements of the To-Be and Radical To-Be scenarios over the As-Is scenario.

Table 5. Differences in Returns on Investment (ROI) of Simulated Scenarios of Repair of Army's Up Armor HMMWV Fleet

		--------------Scenario--------------				
		Variance from As-Was			Variance from As-Is	
No.	Process	As-Is	To-Be	Radical To-Be	To-Be	Radical To-Be
1	Diagnosis at forward station	NA	NA	NA	NA	NA
2	Repair at forward station	NA	NA	NA	NA	NA
3	Retrograde forward station to US depot	-1%	-6%	-18%	-6%	-18%
4	Retrograde forward station to in-theater	NA	NA	NA	-143%	-439%
5	Diagnosis at US depot	6%	336%	NA	331%	NA
6	Repair at US depot	6%	223%	NA	217%	NA
7	Overhaul at US depot	1%	8%	-23%	7%	-24%
8	Transport from US depot to forward	-1%	-6%	-18%	-6%	-18%
9	Diagnosis at in-theater facility	NA	NA	NA	227%	1306%
10	Repair at in-theater facility	NA	NA	NA	283%	-26%
11	Overhaul at in-theater facility	NA	NA	NA	6%	-24%
12	Transport in-theater facility to forward station	NA	NA	NA	-140%	-439%
	TOTAL	22%	74%	95%	51%	73%

Note. NA=*Not applicable* because the process is not used in the scenario.

The positive variances in the bottom row of Table 5 indicate that the advanced technologies significantly improve equipment repair. More specifically, ROI increases 95% from the traditional processes (As-Was) to the envisioned scenario (Radical To-Be) and 73% from the current processes (As-Is) to the envisioned scenario (Radical To-Be). Table 5 also shows losses for shipping equipment back to U.S. depots and back (processes #3, #4, and #12) as the three advanced technologies are increasingly adopted and used (moving right across the rows). This shows that the in-theater and forward station repairs allowed and facilitated by the three advanced technologies make returning equipment to the United States for repairs less attractive with advanced technologies.

Estimating Cost Savings in Shrinking the Mountain

The definition of Return on Investment (ROI), the benefits, and Returns on Investment (Table 4) were used to estimate the costs of each scenario in millions of dollars.[3] Benefits were estimated as the value of the up armor HMMWV fleet, specifically as 23,800 vehicles * $169,428/vehicle[4]= $4,032,386,400. Results are shown in Table 6.

Table 6. Estimated Costs and Savings in Army's Up Armor HMMWV Fleet of Four Scenarios

Sceanrio	Cost ($Mil)	Savings vs. As-Was ($Mil)	Savings vs. As-Was (% fleet value)	Savings vs. As-Is ($Mil)	Savings vs. As-Is (% fleet value)
As-Was	$5,449.17	NA	NA	NA	NA
As-Is	$4,200.40	**$1,248.77**	**31%**	NA	NA
To-Be	$2,724.59	**$2,724.59**	**68%**	**$1,475.82**	**37%**
Radical To-Be	$2,386.03	**$3,063.14**	**76%**	**$1,814.38**	**45%**

The savings shown in Table 6 are consistent with, or conservative, when compared to the results reported by industry adopters of these technologies described previously in this report (e.g., >30% cost savings for 3DST alone and up to 80% for AM). The results suggest that the adoption of the current processes have saved almost $1.2 billion in the up armor HMMWV fleet over the traditional approach and that the additional adoption and use of the advanced technologies can save an additional $1.8 billion or more.

Potential savings of full implementation of an advanced technology strategy (Radical To-Be scenario) for multiple fleets can be estimated using the 45% of fleet value savings in Table 6. Accurate and consistent estimates of the value of U.S. Army equipment are difficult to obtain. However, order of magnitude savings can be estimated using available values. Banian (2013) estimated the value of U.S. Army equipment in Afghanistan to be $28.454 billion, and Cruz (2013) estimated the value of equipment in Afghanistan at the beginning of 2013 as $28 billion. In 2008, the GAO (2008) estimated that the $15.5 billion of DoD materiel and equipment in Operation Iraqi Freedom is theater provided equipment that represents 80% of the total used in Iraq. These estimates suggest a materiel and equipment value of at least $47 billion (28.254+(15.5/.80)=47.7b) for the two operations. Potential savings for future operations of similar scale using the Radical To-Be savings estimate are $21.46b (=$47.7b * 45%). This estimate is based on a single fleet of vehicles. Savings could be larger because multiple fleets of equipment could share repair resources, such as hardware, software, and people, thereby reducing costs further.

---

[3] ROI = *(Benefits–Costs)/Costs, which can alternatively be written as Cost=Benefits/(ROI + 1).*

[4] Cost estimates of a single up armored HMMWV range from $169,248 (DoD, 2014) to $220,000 (Keyes, 2011).

Conclusions

Three advanced technologies were examined for their capability to reduce the cost of shrinking the mountain of equipment generated by military operations. Three-dimensional scanning technology, additive manufacturing, and product lifecycle management have evolved far enough to have demonstrated their potential benefits to diagnosis, repair, and overhaul processes. Forecasted evolutions of the technologies based on the literature were used to develop four realistic scenarios of their application to military equipment repair in the past, present, near future, and distant future. These four scenarios were then modeled using the Knowledge Value Added methodology to estimate returns on knowledge and returns on investment using the up armored HMMWV fleet as an example. The results indicate that the advanced technologies benefit repair operations and generate significant savings, especially by performing damage diagnosis in-theater and at forward stations. The results were used to estimate potential savings of more than $1.8 billion for the up armored HMMWV fleet and at least $21 billion for operations similar to the scale of those in Iraq and Afghanistan.

We conclude that to capture the very large potential savings the DoD should accelerate its adoption of 3DST, AM, and PLM for equipment repair. That acceleration should include testing their use for a broader spectrum of applications (e.g., parts types, processes), the expansion of their use in applications that have been demonstrated to provide benefits, and the revision of processes to exploit these technologies (especially reduce shipping to and from distant depots). Doing so will have important impacts on both practice and research. More military operations support will be located closer and at forward stations. Damage diagnosis and repair will occur much faster, be more accurate, and be targeted. Demands on repair operations will be forecasted in real time based on data from embedded sensors that communicate equipment conditions to support units. Research will be needed to understand and develop effective and efficient processes for these new operations. First steps can include research that learns from existing technology applications and applies that knowledge across multiple equipment types, fleets, and services.

Military repair operations will experience growing pains as the adoption of advanced technology force operational and support changes. But these changes will result in very large cost savings and increased operational flexibility. By exploiting advanced technologies, the DoD can accelerate and reduce the cost of shrinking the mountain, increase the value of that materiel, and improve the operational capability of U.S. military forces.

References

Allard, P.-H., Lavoie, J.-A., & Fraser, J.-S. (2013). Improvement of aircraft mechanical damage inspection with advanced 3D imaging technologies. In *Proceedings of Fifth International Symposium on NDT in Aerospace*. Singapore: NDT. Retrieved from http://www.ndt.net/article/aero2013/content/papers/13_Allard.pdf

Anusci, V. (2015). You're in the Army now: 10 future applications for 3D printing. Retrieved from https://all3dp.com/youre-in-the-army-now-10-future-applications-for-3d-printing/

Aquipedia. (2007). Reset of military equipment, primary functional area: Life cycle logistics. Retrieved from https://dap.dau.mil/acquipedia/Pages/ArticleDetails.aspx?aid=1e96e8de-1e6b-493a-b81f-8da6f41aa897

Banian, D. (2013). From hard to harder: Iraq retrograde lessons for Afghanistan. Retrieved from https://www.army.mil/article/112409

Congressional Budget Office (CBO). (2007, September). *Replacing and repairing equipment used in Iraq and Afghanistan: The Army's reset program*. Retrieved from https://www.cbo.gov/sites/default/files/110th.../09-13-armyreset.pdf

Earls, A., & Baya, V. (2014). The road ahead for 3-D printers. Retrieved from http://www.pwc.com/us/en/technology-forecast/2014/3d-printing/features/future-3d-printing.html

GAO. (2008, September). *Operation Iraqi Freedom, actions needed to enhance DOD planning for reposturing of U.S. forces from Iraq*. (GAO-08-930). Retrieved from http://www.gao.gov/products/GAO-08-930

GAO. (2012, May). *Warfighter support: Army has taken steps to improve reset process, but more complete reporting of equipment and future costs is needed*. (GAO-12-133). Retrieved from http://www.gao.gov/products/GAO-12-133

Hendren, J. (2007, February 10). Army struggles with backlog of damaged gear. TEXARKANA, Texas: ABC News. Retrieved from http://abcnews.go.com/WNT/story?id=2865968&page=1

Housel, T., Hom, S., Ford, D., & Mun, J. (2015). Benchmarking naval shipbuilding with 3D laser scanning, additive manufacturing, and collaborative product lifecycle management. In *Proceedings of the 13th Annual Acquisition Research Program*. Monterey, CA: Naval Postgraduate School.

Keyes, C. (2011). Steep cost of military vehicles outlined in Army report. Retrieved from http://www.cnn.com/2011/US/01/27/army.vehicle.costs/

Lu, B., Li, D., & Tian, X. (2015, March). Development trends in additive manufacturing and 3D. *Engineering, 1*(1), 85–89. Retrieved from http://www.sciencedirect.com/science/article/pii/S2095809916300492

Shilovitsky, O. (2016, February) 5 trends that will disrupt the future of product lifecycle management (PLM) [Blog]. Retrieved from http://www.ibmbigdatahub.com/blog/5-trends-will-disrupt-future-product-lifecycle-management-plm

Smith, R. (2015), 5 incredible trends that will shape our 3D printed future. *Forbes*. Retrieved from http://www.forbes.com/sites/ricksmith/2015/07/07/5-incredible-trends-that-will-shape-our-3d-printed-future/#25756d9a707f

Solis, W. (2006). *DEFENSE LOGISTICS–Preliminary observations on equipment reset challenges and issues for the Army and Marine Corps*. (GAO-06-604T). Retrieved from http://www.gao.gov/products/GAO-06-604T

Taktikz. (2017). Sustainment terms. Retrieved from http://taktikz.com/support/frequently-asked-questions/gdi/defense/sustainment/sustainment-terms/

Thomas, D. S., & Gilbert, S.W. (2014, December). *Costs and cost effectiveness of additive manufacturing* (Special Publication [NIST SP) 1176). Retrieved from https://www.nist.gov/node/562081

U.S. military to use 3D modeling to maintain their MRAP Cougars [Blog]. (2014, June). Retrieved from http://www.smartgeometrics.com/blog/engineering/us-military-to-use-3d-modeling-to-maintain-their-mrap-cougars/

Acknowledgments

The authors thank Bernice Pettinato for literature search assistance and the Acquisition Research Program for financial support of this work.

Using Additive Manufacturing to Mitigate the Risks of Limited Key Ship Components of the Zumwalt-Class Destroyer

LT Xiao Y. Wang, USN

LCDR James R. Whitworth, USN

Introduction

The purpose of this project was to explore the benefits of using a combination of additive manufacturing (AM), performance-based logistics (PBL), and open systems architecture (OSA) to mitigate the risks of limited key ship components for the Zumwalt-class destroyer (DDG 1000) program. Specifically, this project was focused on current industry's capability for AM and the implementation of AM in the near future. Research was conducted in three phases. First, this research reviewed the problems and challenges within the defense industry. Next, this research reviewed the previous research on intellectual property (IP) concerns with AM (particularly, insourcing versus outsourcing) and the latest AM applications in the marketplace and defense industry. Finally, this research focused on DDG 1000 program documents, including the Acquisition Strategy (AS), the Life-Cycle Sustainment Plan (LCSP), and a Diminishing Manufacturing Sources and Material Shortages (DMSMS) analysis. By conducting a comparison of DDG 51 and DDG 1000 and analyzing an AM arrangement among Airbus, Systemanalyse and Programmentwicklung (SAP), and United Parcel Service (UPS), this research concludes that the government can use AM, with a properly structured PBL arrangement and OSA, to substantially mitigate risks, lower operation and support (O&S) costs, and effectively improve system readiness.

Zumwalt-class destroyer (DDG 1000) is a three-ship program that represents the pinnacle of state-of-the-art technology. Because of technologies, intellectual properties, and scale economies, DDG 1000 is in a sole-source, or limited sources, acquisition environment. The risks associated with a limited supplier base could threaten the part support on many key ship components and the overall performance of its service life for the next 25 years or more. For cost saving purposes, all three ships will have a homeport in San Diego, CA, where organic repair, off-ship maintenance, and performance-based logistic support take place. The DDG 1000 program is also facing budget cuts, program cost growth, and competition from other classes of ships; therefore, Program Executive Office (PEO) Ships and the DDG 1000 program office must find ways to mitigate the risks of key ship components and enhance system performance with a sound life cycle sustainment strategy.

Traditional approaches for operating and maintenance are accomplished with organic repair capabilities or contracted services. Due to the technology complexity and existing organic capabilities, a combination of organic support and performance-based logistics (PBL) has been identified as part of DDG 1000's life cycle sustainment plan. Regardless of the approaches, either the government or the chosen PBL providers will have to tackle the obsolescence issues and address the issues associated with a limited supplier base. Traditionally, the decision-maker will have to decide on either a lifetime-buy or bridge-buy decision, based on industry data and the obsolescence management forecast, and anticipate failure rates to ensure that the needed parts are available for the operation and support of the systems. The advent of additive manufacturing (AM) and recent technology advancement can eliminate the need for a lifetime or bridge-buy decision, reduce ship's operating and maintenance costs, and enhance system performance. Research on AM developments is used to identify capability gaps and explore opportunities for improving system readiness.

In order to introduce AM as part of the solution, this project first examined the benefits and limitation of PBL and assessed the competition requirement for federal acquisition strategy and the challenges in obsolescence management. This project then verified that PBL and OSA are part of DDG 1000's acquisition strategy, as they are the prerequisites for entering a contractual agreement with contracted service providers for Operation and Support (O&S) and enabling system interoperability. This project subsequently compared operating and support characteristics between Arleigh Burke–class (DDG 51) and Zumwalt-class (DDG 1000) ships and assessed the ability of the DoD to expand DDG 1000's logistic support footprint, similar to the arrangement among Airbus, Systemanalyse and Programmentwicklung (SAP), and UPS.

The purpose of the study was to research the latest AM developments within the commercial marketplace and defense industry and explore the ways that AM can help to drastically reduce the risks of limited key ship components. The project answered the following questions:

- How should the government structure PBL contracts that will incentivize the use of AM?
- If the government decides to insource, what are the considerations in make-or-buy decisions?
- How can the DDG 1000 program leverage the capabilities of AM for its existing and future requirements?

Primary research data was provided by the Zumwalt-class Program Office (PMS 500). Secondary research was collected from public resources. Based on the findings of this research, it is imperative to have AM, properly structured PBL arrangements, and well thought-out OSA strengthen each other and mitigate the risks of limited key ship components that are associated with their supplier base. Naval Undersea Warfare Center (NUWC), PMS 500, and contractors could jointly identify parts as candidates for AM solutions. PMS 500 should also engage other DoD agencies on AM capacities and request information from defense contractors on their planned use of AM capabilities for part support.

Finding 1 and Recommendations

In order to take advantage of the capabilities and potential that AM offers, the government needs to structure performance-based arrangements that will help to extract innovation, motivation, and collaboration from its contractors. As the 2016 update to the DoD PBL guidebook stated, "PBL is not a one-size fits all tool … evidence provides a compelling case that performance-based sustainment is both a successful and robust strategy" (ASD[L&MR], 2016). While PBL arrangements can transfer the risk of managing O&S to a contractor to a certain degree and insourcing can provide some assurance, PBL and insourcing can still have shortfalls.

There are additional challenges for entering a PBL relationship with commercial vendors and defense contractors with the aim of taking advantage of the most revolutionary manufacturing process. First, the government has to incorporate AM as part of its requirement, acquisition strategy, and sustainment plan. Then the government needs to solicit the ideas, offers, and solutions from the marketplace and defense industry. Since the current state of technology makes AM more ideal for low-volume and low-quantity production, the government will need to use incentives to elicit desired behaviors and extract performance outcomes. It is not as simple as increasing the profit margin and reimbursing allowable costs, but entails careful planning, analyzing, monitoring, and evaluating with adequate goals, metrics, and methodologies.

The government understands its requirements well, or at least, the quantity and the delivery schedule of the requirements, while the contractors know their capabilities better. If the government does not clearly specify the requirement of incorporating AM as part of the PBL contract, the contractor will have the ambiguity and freedom to decide how to satisfy the government's needs. Since traditional manufacturing facilities are still in use and a large order quantity provides good profit, some contractors will have less incentive to introduce AM as part of a solution and will continue to expect that a large quantity was the primary means to achieving cost savings. More importantly, it is in the government's best interest to identify AM vendors through market research as soon as practical and promote competition.

While the 2016 release of the PBL guidebook provides the DoD with the 10 tenets for PBL arrangement, the strategic considerations for IP rights, the data collection phase process, and the steps for the implementation of PBL, this project looked into the FAR and DFARS, which identified two specific incentives the government can use to elicit the desired outcomes. AM requires companies to sink substantial investments, in both resources and manpower, to keep pace with the new technologies/sub-technologies that are constantly evolving in the marketplace. At the same time, existing, traditional manufacturing facilities, resources, and processes are competing for resources and remain significant for their known advantages. Therefore, it is important to use the efficiency factor during PBL contract negotiation to spur investment in innovative manufacturing processes, particularly in AM. Also, this project shows that most of the defense contractors who aggressively pursue AM are the well-established industry giants. In order to maximize small business participation in AM and extract nontraditional solutions, the government needs to use the FCCOM to assist, reimburse, and compensate the contractors' capital investment.

As the government uses significant, irresistible incentives to lead the industrial revolution in AM, the government can effectively reshape the defense industry landscape. The government can reduce its risk of a limited supplier base by having a second supplier, or multiple suppliers, for its system acquisition and service support, particularly in the low-volume, low-quantity defense articles. In short, the PBL guidebook laid the generic, strategic framework for contract support, and this project identified the actionable items for execution by evaluating the environment, requirement, and characteristics of AM. A properly structured PBL contract could help to alleviate the workload of the contracting officer by placing this requirement on the prime contractor, thus helping to improve competition and achieve cost savings. With an adequate, carefully designed, and innovative PBL approach, the government can encourage research and development AM while helping contractors to improve the quality, reliability, and performance of their products and services.

For the past eight years, there has been an increasing push towards government insourcing. Insourcing, indeed, is a way to mitigate risk and provide some assurances when the marketplace cannot satisfy government requirements. However, reliance solely on insourcing could hurt the defense industry by eliminating the need for some companies.

Finding 2 and Recommendations

PBL is more of a buzzword than an attainable goal if the steps identified in the PBL guidebook and requirements from customers are not achievable or attainable. PBL is also not a one-size-fits-all tool (ASD[L&MR], 2016). The same rule applies here: Insourcing is not the ultimate solution. As this project shows in its research of the obsolescence management case associated with the Parasense sensor, even with dedicated government teams for obsolescence management, engineering advisories for risk monitoring and mitigation, organic capabilities for repairing, and a procurement contract and commercial supplier in place, insourcing does not address all the challenges and complexity of supporting complex defense systems. Moreover, the government does have certain limitations and constraints for ensuring that parts meet specifications. Despite the fact that the DoD has engaged AM for more than 20 years, the government has been a little behind on establishing the qualification or certification process for the use of 3D-printed parts as critical items on weapon systems. In contrast, the qualification, certification, and standardization of AM parts to meet FAA requirements and fielding for commercial use came from GE, without the direction or requirement from any government entity. NAVAIR's recent effort to shorten the AM parts certification process to weeks or even days by developing industry standards for 3D-printed parts is probably the area in which the government can most effectively add value to its operation and make sure it has access to this required capability.

As a result, the government's main challenge is to find the right balance between its essential need for insourcing and the many benefits AM capabilities offer. To do so, the government needs to evaluate its capacities and mission profiles carefully. First and foremost, the government is not in the manufacturing business and should focus on the inherent government functions that cannot be contracted out. In the operation and support of defense weapon systems, many parts and services are considered to be mission critical; however, producing these parts and providing the maintenance and services for them are not inherently government functions. Secondly, parts and services can definitely lead to a life-or-death issue, especially on the battlefield or in contingency situations. However, it is more important that the government can manage and satisfy its requirements through the proper sourcing strategies and channels, instead of providing the services or materials in-house.

To be more specific, perhaps the government needs to find ways to manage the IP rights for the use of AM and provide the regulatory oversight on the standardization, qualification, and certification of AM parts. Instead of relying on contractors to tell the government what to buy, how to manage, and how much to pay, the government probably should focus on insourcing those inherent government functions and be able to coordinate its efforts in the use of AM. Precisely as the SECNAV stated in his memo to the CNO, CMC, and ASN(RD&A), the DoN needs to increase the development and integration of AM, as well as develop the ability to qualify and certify AM parts (SECNAV, 2015). Moreover, the second half of SECNAV's 2015 memo identified standardization of the digital AM framework, end process integration, establishment of the DoN advanced integrated digital manufacturing grid, and formalized access to AM education, training, and certifications for the DoN workforce as more important than organic capabilities. Through the evaluation of the arrangement among Airbus, SAP, and UPS, this research project showed that Airbus is more concerned with selecting the right data management firm, SAP, and capable AM manufacturer, UPS, to satisfy the requirements for meeting its operational and logistical support demands for Airbus's global network.

Finding 3 and Recommendations

AM can improve competition and lower the risks associated with a limited supplier base by adding a second competitor, lowering the nonrecurring costs, eliminating the need for an economy order quantity, and achieving cost savings. AM could allow rapid prototyping—with an OSA design, more small businesses can research, develop, and test their products as subcontractors and help to improve the DDG 1000's capabilities, reliabilities, and sustainability. Last but not least, the use of AM will allow easier incorporation of the open architecture idea to develop new systems while using interface management.

Southard's (2016) iPDA analysis showed that the little bits and pieces of a circuit card can significantly affect the DDG 1000's mission and the obsolescence management forecast is only as good as the current data provided by the marketplace and industry for the next five years. At the same time, the obsolescence case of the Parasense sensor showed that the government sometimes needs to forecast further out into the future, perhaps 25 years or more. In the event that AM becomes the predominant manufacturing process for many low-demand, low-volume parts, the government can take advantage of this revolutionary market dynamic with proper planning. By using a properly structured approach with PBL, OSA, and AM, the government can predict, anticipate, and manage the risks of limited key ship components.

For existing requirements, the government could look into existing AM efforts among government agencies and leverage the equipment on hand from the multitude of entities that have already embraced this technology, including the Department of Energy, NASA, the USS *Essex*, Marine Corps Air Station (MCAS) Cherry Point, and Walter Reed National Military Medical Center in Bethesda, MD (SECNAV, 2015). NUWC, or a designated team, could look into the organic AM capabilities and determine if the government can 3D-print some iPDA parts. Moreover, since GE is not only the maker of the LEAP engine for commercial aviation but also the manufacturer of the LM-2500 engine for the Navy's Arleigh Burke–class destroyers, Ticonderoga–class cruisers, and America–class amphibious ships, the government should investigate the capabilities that GE has for certain existing engines' parts support and system upgrades. Last but not least, the DDG 1000 program can bring the composite deckhouse back onto the negotiation table since 3D-printed composite materials were approved for structural components of commercial ships in 2014 (Job, 2015).

For future requirements, the government should incorporate and insist on the use of new technologies to produce obsolete and low-volume/demand requirements when negotiating PBL arrangements with contractors. Since BAE is currently the prime contractor for DDG 1000's combat system suite and is also aggressively pursuing AM capabilities, it is important to have a discussion on the use of AM for system acquisition and O&S planning. BAE is also more likely to become the PBL provider based on current trends and the defense industry environment the government is in; therefore, it is imperative for the government to understand and define its requirement to develop and negotiate the proper measurement metrics for program execution. For certain military requirements, such as the DDG 1000's propeller, the challenge is to meet those stringent standards under extreme conditions. For example, a 3D-printed propeller for a naval warship would need to pass the shock test and sustain a prolonged period of high-speed maneuvering. It will be worth the effort to investigate the EBAM metal 3D printer, located at Lockheed Martin's manufacturing facility, which is capable of metal-printing parts up to 19 feet long. The government should direct its effort towards AM and take advantage of the SECDEF's plan to spend $72 billion on R&D (Buren, 2016).

Similar to PBL and insourcing, AM is a revolutionary manufacturing process that has certain limitations with the current state of AM technologies. AM is also not a one-size-fits-all tool/solution and might not be cost effective for every application. For example, manufacturers will continue to use their existing facilities and resources to produce those high-volume and low-complexity parts until the costs of maintaining those resources are no longer economically sound. In short, AM or 3D printing by themselves may not be the solution for many existing and future requirements; however, when the government can combine them with PBL and an open system approach, the government can significantly lower a program's cost, performance, and schedule risks. It is predictable that the Navy, especially the DDG 1000 program office, can benefit from AM development immediately and in the near future.

Conclusions

The DDG 1000 program's re-baseline, budget cuts, technology maturity issues, cost increases, and other unknown risks could have led these three ships into a perfect storm and a much cloudier, muddier future. Additionally, most of the DDG 1000 suppliers are sole-source and therefore enjoy a certain monopoly of power in the marketplace. Even though DDG 1000's increased parts costs associated with a sole-source or limited sources environment is a valid concern, the fact that the mere existence of DDG 1000's suppliers can significantly affect the program's performance is the primary concern for long-term O&S planning. As this project shows, without an effective way to lower startup costs and extract ROI, the government will not obtain competition or maintain a healthy industry base. AM, a revolutionary manufacturing process, is the potential answer to these problems.

The DDG 1000's OSA and PBL, by design or by accident, have jointly crafted an environment for the introduction and implementation of AM. The DDG 1000 program does not have scale economies due to the number of ships; therefore, DDG 1000 needs to seriously consider AM as a means to satisfy its low-volume, infrequent-demand requirements, as well as to mitigate the risk of a limited supplier base. AM can help to improve material availability and alter the traditional obsolescence management approach that is more likely to result in lifetime buy decisions with possible limitations. With PBL, OSA, and AM, the program office can invite more interested parties to participate and thus mitigate the risk of losing existing contractors. Since many of the major AM developers—such as SAP, BAE, Lockheed Martin, and Raytheon—are also the primary service providers of the DoN, and since the Navy has similar or better resources for DDG 1000 to mirror the agreement that Airbus, SAP, and UPS developed, AM is a viable solution in mitigating the risks identified in this research.

AM will not replace conventional manufacturing methods for high-volume, low-complexity parts in the near term and foreseeable future; however, AM will continue to evolve as the process matures and will significantly alter make-or-buy decisions for low-volume, low-quantity items. DDG 1000, as the tech engine for the fleet, will reap the benefits of rapid prototyping, faster production, better quality parts, lower prices, and minimum risks that AM offers.

Summary

Looking forward, the advent of AM and associated future technology advancements will continue to reshape the industry landscape and challenge the business decision-making process. From the commercial marketplace to the private defense industry, AM is aggressively pursued and incorporated into business decisions. The challenge for the DoD acquisition community, across the spectrum from system engineering to contracting, is to incorporate AM in decision-making throughout all phases of the product life cycle. There are many uncharted areas for the use of AM developments to identify the capability gaps, to improve system readiness, and to meet future mission requirements; therefore, the DoD must lead from the forefront and take a holistic approach to integrating AM. Defense system acquisition, like the DDG 1000 program, can significantly benefit from the use of AM and drastically reduce the risks of limited key components.

References

Assistant Secretary of Defense for Logistics and Materiel Readiness (ASD[L&MR]). (2016). *PBL guidebook: A guide to developing performance-based arrangements.* Washington, DC: DoD.

Buren, A. (2016). *US Navy plans to fly a Boeing V-22 Osprey aircraft with 3D printed part in June*. Retrieved from http://www.3ders.org/articles/20160518-us-navy-announces-test-flight-with-a-boeing-v-22-with-3d-printed-part-in-june.html

Job, S. (2015). Composite hatch cover approved for use in commercial ship. *Reinforced Plastics, 59*(2), 58.

Secretary of the Navy (SECNAV). (2015, September 3). *Additive manufacturing/3-D printin*g [Memorandum]. Washington, DC: Author.

Southard, C. (2016). *DMSMS analysis of iPDA circuit card.* Keyport, WA: Navy Undersea Warfare Center (NUWC).

Panel 15. Informing Future Ship Design Projects

Thursday, April 27, 2017	
1:45 p.m. – 3:15 p.m.	**Chair: Rear Admiral William J. Galinis, USN**, Program Executive Officer, Ships
	Flexible and Adaptable Ship Options: Assessing the Future Value of Incorporating Flexible Ships Design Features Into New Navy Ship Concepts
	Johnathan C. Mun, Naval Postgraduate School Thomas J. Housel, Naval Postgraduate School
	Persistent Platforms—The DDG 51 Case
	Ira Lewis, Naval Postgraduate School
	Applying Principles of Set-Based Design to Improve Ship Acquisition
	Eric Rebentisch, Massachusetts Institute of Technology Commander John A. Genta, USN

Rear Admiral William Galinis, USN—is a native of Delray Beach, FL. He is a 1983 graduate of the U.S. Naval Academy where he received a Bachelor of Science in Electrical Engineering. He holds a Master of Science in Electrical Engineering from the Naval Postgraduate School.

Galinis' tours as a surface warfare officer included damage control assistant aboard USS *Vreeland* (FF 1068) and engineer officer aboard USS *Roark* (FF 1053). He was selected for transfer to the engineering duty officer community in September 1991.

Galinis' initial engineering duty tour was with the supervisor of Shipbuilding, Conversion, and Repair, New Orleans, where he worked on both new construction and repair projects including assignment as the PMS 377 program manager's representative for the LSD (CV) Shipbuilding Program. He subsequently served as the senior damage control inspector for the Board of Inspection and Survey, Surface Trials Board, as well as in a number of program office and staff positions including the DD 21 and LPD 17 Program Offices, Office of the Chief of Naval Operations in the Requirements & Assessments Directorate (N81), and in the Office of the Deputy Assistant Secretary of the Navy for Shipbuilding as the chief of staff.

His command assignments included LPD 17 program manager—leading the commissioning of the first four ships of the LPD 17 San Antonio Class, delivering the fifth ship and starting construction on four additional ships; supervisor of shipbuilding, Gulf Coast overseeing Navy ship construction projects and Foreign Military Sales work in shipyards along the Gulf Coast and Wisconsin; and as the commanding officer of the Norfolk Ship Support Activity (NSSA) where he led ship maintenance and repair efforts.

Galinis' first flag assignment was deputy commander for surface warfare, Naval Sea Systems Command (NAVSEA 21)/commander, Navy Regional Maintenance Center, responsible for managing critical modernization, maintenance, training, Foreign Military support contracts, and inactivation programs.

Currently, Galinis is serving as program executive officer, Ships, where he is responsible for Navy shipbuilding for surface combatants, amphibious ships, logistics support ships, support craft, and related foreign military sales.

Galinis has received various personal, unit, and service awards including three Navy Battle "E" awards.

Flexible and Adaptable Ship Options: Assessing the Future Value of Incorporating Flexible Ships Design Features Into New Navy Ship Concepts

Johnathan C. Mun—is a research professor at NPS and teaches executive seminars in quantitative risk analysis, decision sciences, real options, simulation, portfolio optimization, and other related concepts. He received his PhD in finance and economics from Lehigh University. He is considered a leading world expert on risk analysis and real options analysis. Dr. Mun has authored 12 books and is the founder and CEO of Real Options Valuation Inc. [jcmun@realoptionsvaluation.com]

Thomas J. Housel—specializes in valuing intellectual capital, knowledge management, telecommunications, information technology, value-based business process reengineering, and knowledge value measurement in profit and non-profit organizations. He is a tenured full professor for the Information Sciences (Systems) Department at NPS. He has conducted over 80 knowledge value added (KVA) projects within the non-profit, Department of Defense (DoD) sector for the Army, Navy, and Marines. Dr. Housel also completed over 100 KVA projects in the private sector. The results of these projects provided substantial performance improvement strategies and tactics for core processes throughout DoD organizations and private sector companies. [tjhousel@nps.edu]

Abstract

To successfully implement the Surface Navy's Flexible Ships concept, PEO-SHIPS requires a new methodology that assesses the total future value of various combinations of Flexible Ships design features and how they will enable affordable warfighting value over the ship's full service life. Examples of Flexible Ships design features include decoupling payloads from platforms, standardizing platform-to-payload interfaces, allowance for rapid reconfiguration of onboard electronics and weapons systems, preplanned access routes for mission bays and mission decks, and allowance for sufficient growth margins for various distributed systems. The current research analyzes the application of strategic Real Options Valuation methodology within the Integrated Risk Management process to assess the total future value of Flexible Ship design feature options. This approach can be used to support the Future Surface Combatant Analysis of Alternatives technique. The current research has the explicit goal of proposing a reusable, extensible, adaptable, and comprehensive advanced analytical modeling process. This methodology is designed to help U.S. Navy decision-makers in quantifying, modeling, valuing, and optimizing a set of ship design options to support a business case for making strategic acquisition decisions in the context of various quantifiable uncertainties.

Introduction

The U.S. Navy is tasked with fulfilling its missions globally in environments with rapidly changing threats using an equally rapidly evolving technological base of ship platform, mission, electronic, and weapon systems. The challenge the U.S. Navy faces is to retain and maintain sufficient military capabilities during wartime as well as a deterrent during peacetime, with the added goal of minimizing major intrusive, time-consuming, and costly modernization throughout a ship's service life by incorporating Modular Adaptable Ships (MAS) and Flexible and Adaptable Ship Options (FASO) in the ship design. Pursuing this goal has the added benefit of allowing the Navy to affordably and quickly transform a ship's mission systems over its service life to ensure it maintains its required military capabilities (Doerry, 2012).

Historically, naval ship design includes robust fixed structural features that limit the ability to include options for any future capabilities that may require design changes. For instance, any major requirement changes needed to meet critical operational tasks during wartime would necessitate a major modernization effort or decommissioning the existing ship prior to the end of its service life and replacing it with a newly commissioned ship. The concept of MAS and FASO, if applied correctly, with the optimal options implemented, would reduce the need for costly and lengthy major mid-service-life intrusive modernizations, as well as increase the existing platform's flexibility to adapt to new requirements utilizing a faster and cheaper alternative.

The concept of FASO is not new to the Navy. In fact, benefits of MAS/FASO concepts have previously been detailed by Jolliff (1974), Simmons (1975), Drewry (1975), and others. Even as recently as 2015, the Naval Sea Systems Command's (NAVSEA) PEO SHIPS released a presentation on Flexible Ships, detailing its *Affordable Relevance Over the Ship's Life Cycle* (Sturtevant, 2015). In it, Director of Science and Technology, Glen Sturtevant, noted that the current and future main challenges confronting Surface Navy include unknown but evolving global threats while having to manage an accelerated pace of technological changes, coupled with rising costs and declining budgets. The analysis found that ships currently cost too much to build and sustain; the ships (Platforms) are too tightly coupled with their capabilities (Payloads); and inflexible and fixed architectures of legacy ships limit growth and capability upgrades or result in lengthy and costly upgrades. The effects of these issues, of course, are compounded by ever-evolving global threats that introduce significant uncertainty.

ADM Greenert (Chief of Naval Operations) and VADM Rowden (Commander of Naval Surface Forces) in past speeches echoed the idea that the ability to quickly change payloads and having modularity on ships would maximize the service life of ships and allow faster and more affordable upgrades to combat systems and equipment.

Some examples of MAS and FASO that had been espoused in Navy ship design research literature, such as in Sturtevant (2015), Doerry (2012), Koenig (2009), Koenig, Czapiewski, and Hootman (2008), and others, include Decoupling of Payloads from Platforms, Standardizing Platform-to-Payload Interfaces, Rapid Reconfiguration, Preplanned Access Routes, and Sufficient Service Life Allowance for Growth. These FASO areas can be applied to a whole host of systems such as weapons, sensors, aircraft, unmanned vehicles, combat systems, C4I, flexible infrastructure, flexible mission bays and mission decks, vertical launch systems (VLS) for various multiple missile types, future high-powered surface weapons (laser weapon systems and electromagnetic railguns), and modular payloads (e.g., anti-submarine warfare, special operations, mine warfare, intelligence gathering, close-in weapon systems, harpoon launchers, rigid hull inflatable boats, gun systems, etc.).

The concepts of Adaptability and Flexibility (plug-and-play concepts of rapidly removing and replacing mission systems and equipment pier-side or at sea), Modularity (common design interface and modular components that will greatly simplify adding, adapting, modifying, or modernizing a ship's capabilities), and Commonality/Scalability (capabilities that are built independently of a ship by using standardized design specifications that allow similar systems to be placed across multiple ship platforms) are concepts that can take advantage of the capabilities of strategic Real Options Valuation (ROV) analytical methodologies. ROV has been used in a variety of settings in Navy ship maintenance, signal intelligence, and shipbuilding contexts, as well as being widely used in various industries including pharmaceutical drug development, oil and gas exploration and production, manufacturing, start-up valuation, venture capital investment, information technology infrastructure, research and development, mergers and acquisitions, intangible asset valuation, and others. The current project applies the same flexibility modeling empowered by Real Options Valuation methods to identify the optimal ship design alternatives.

This current research acknowledges that the U.S. Navy has attempted to incorporate FASO and MAS capabilities in its ship design of Future Surface Combatants (FSC). Further, the Navy acknowledges that there is significant value in terms of being able to rapidly upgrade FASO ships at a lower cost, while extending the ships' service life, all the while being able to quickly adapt to changes in both external threats and internal new technologies. As such, this current research is not meant to identify said FASO/MAS platforms or payloads per se. Rather, this research is designed to examine previously identified platforms such as the DDG 51 Flight III where there are opportunities to insert flexible ship features. This analysis is limited to the context of the domain of Anti-Submarine Warfare (AWS).

The current research reviews a series of quantitative tools and techniques that may be useful in developing a Business Model or Business Case Analysis to support strategic decision-making, under uncertainty. Specifically, it will identify, model, value, and optimize the various strategic real options in flexible ship designs. Currently, there is only a limited set of real-life applications of FASO/MAS in ship design and they are classified; therefore, actual empirical data is not used in this research. In addition, because the objective of this research is to illustrate, in detail, a potential business case modeling process and analytical methodology (such that the method and process can be replicated and used in all future FASO/MAS design decisions), subject matter expert (SME) opinions, publicly available information, and certain basic assumptions or rough order of magnitude (ROM) estimates were used. The use of said ROM or SME inputs does not detract from the analytical power, efficacy, or applicability of these methods.

In summary, the process will:

- Identify which FASO/MAS options have a positive return on investment (i.e., the benefits outweigh the costs).
- Model Uncertainty and Risks (i.e., Monte Carlo Risk Simulations will be applied to simulate hundreds of thousands of possible scenarios and outcomes to model the volatility and ever-changing global threat matrix).
- Frame and Value the Ship Design Options (i.e., each design option will be vetted and modeled, where the options will be framed in context and valued using cost savings for rapid upgrades at lower costs, costs to design and implement these FASO/MAS options, and estimate potential military value using the Knowledge Value Added method to monetize expected military value).
- Optimize the Portfolio of Options (i.e., in the context of a given a set of FASO/MAS design options with different costs, benefits, capabilities, and uncertainties, given constraints in budget, schedule, and requirements).

The Real Options Solution in a Nutshell

Simply defined, real options is a systematic approach and integrated approach for valuing real options (e.g., ship design options). This approach employs financial theory, economic analysis, management science, decision sciences, statistics, and econometric modeling in applying options theory to value real physical assets (e.g., options for weapons insertion on a ship platform). It is not designed for valuing financial options (e.g., calls, puts in the stock market context). It is appropriate for a dynamic and uncertain environment where decisions have the possibility of being flexible in the context of strategic acquisition investments by, valuing investment opportunities, and projecting required capital expenditures. Real options are crucial in:

- Identifying different acquisition investment decision pathways or projects that management can navigate given highly uncertain conditions (e.g., new enemies, new technologies).
- Valuing each of the strategic decision pathways and what they represent in terms of financial viability and feasibility.
- Prioritizing these pathways or projects based on a series of qualitative and quantitative metrics.
- Optimizing the value of strategic investment decisions by evaluating different decision paths under certain conditions or using a different sequence of pathways that can lead to the optimal acquisition strategy.
- Timing the effective execution of acquisition investments and finding the optimal trigger values and cost or value drivers.
- Managing existing or developing new optionalities and strategic decision pathways for future opportunities (e.g., technology or weapon insertions).

ROV is useful for valuing a project, alternative path, implementation option, or ship design through its strategic options especially in capital-intensive investment decisions under uncertainty. In a traditional cost-benefit and cash flow model, the ROI or cost-benefit question, is presumed to be an accurate, but static, representation of the potential future value of acquisition options. In fact, some of the answers generated through the use of traditional cash flow models are flawed because the model assumes a static, one-time decision-making process with no recourse to choose other pathways or options in the future. In contrast, the real options approach takes into consideration the strategic managerial acquisition options certain projects create, under uncertainty, and the decision-makers' flexibility in exercising or abandoning these ship design options at different points in time, when the quantitative level of uncertainty has decreased or has become known over time.

Real Options Valuation Applications in the U.S. Department of Defense

This section provides a quick snapshot of the various ROV option types and their relevance to the DoD in general, as well as applications within the scope of the current research.

Option to Wait and Defer (Ability to Wait Before Executing)

An option to defer allows the holder the option, but not the obligation, to execute a certain strategy when situations make it optimal to do so. An option to wait and defer provides the holder with the following advantages:

- A portfolio of capabilities and readiness for immediate deployment can be created and maintained with the use of options to defer. If the predeveloped payload or platform options exist, they will allow rapid change out of equipment and integration of new weapons or electronics systems, without the excessive schedule and cost penalties.
- Options to defer allow ship designers to incorporate modernization and upgrade options into the ship design early on, and to defer the exact configuration of the ship until a future date when uncertainties on capability requirements are resolved over the passage of time (midlife of the ship's lifespan), actions (new missions), and events (wartime, peacetime).

- By creating design options and design flexibility specifically for mission and weapon systems that are anticipated to have the maximum change over the lifespan of a ship, and at the same time, using common bow and stern configurations, any changes in future capability requirements can be accommodated quickly and cheaply.
- Other applications within the DoD include the following:
    - Build or Buy Options (Buy versus Lease Options). That is, should a technology be developed internally, or should commercially available off-the-shelf applications be used?
    - Multiple Contracts and Vendors. Having multiple vendors or contracts in place that may or not be executed increases the chances of corporate survivorship and an existing military industrial base to ensure future uncertain demands are met.
    - Capitalizing on other opportunities while reducing large-scale implementation risks, and determining the value of P3I and R&D (parallel implementation of alternatives while waiting on technical success of the main project, and no need to delay the project because of one bad component in the project).
    - Low-rate initial production (LRIP), advanced weapons R&D, advanced technology demonstrations, and weapons and systems prototyping. Provide the right of first refusal to test and see the results (deferring the final decision) until the outcomes of said trials are evident.
    - There is significant Value of Information in forecasting cost inputs, capability requirements, schedule risks, and other key decision metrics by deferring decisions until a later date, but having the option ready to be triggered at a moment's notice.
    - Military intervention strategies include the naval option, the air option, go-long versus go-deep versus go-home option, first strike option, surge option, force mix option, and deterrent options.

Option to Switch (Ability to Switch Applications)

An option to switch allows the holder the right, but not the requirement or obligation, to maintain the current status quo or to switch among a variety of predetermined options. The decision on which option to execute is deferred until a future date when exact needs and specifications are known, and the optimal option is then executed.

- Standardization and Modularity. By incorporating options to ensure ISO standards for containers, tie-down systems, mission bays, and support structures, ships can take on multi-mission payloads quickly and efficiently.
- Flexible infrastructure options within a ship, such as open power, open HVAC, open data cabling, open outfitting, and open structure, allow ships to be adaptable and reconfigured for different missions quickly without major rework such as stripping and welding.
- Other applications within the DoD include, but are not limited to:
    - Switching vendors in Open Architecture (OA) and modular concepts allows the U.S. Navy to use multiple vendors for similar parts, ensuring healthy price and quality competition sustainment in the industry, as well as existing parts suppliers for the future.

- Readiness and capability risk mitigation can be obtained through ensuring multiple vendors and a strong military industrial base.

Simultaneous Compound Option (Parallel Development)

Simultaneous and parallel development efforts are sometimes used to reduce critical path and schedule risks. The risk of technical failures during development or schedule delays, especially when speed is critical, are mitigated with this simultaneous option where multiple systems are designed in parallel.

- By designing multiple payloads (combat subsystems or electronic subsystems) in parallel with the platform (ship design), newer weapons systems may be ready for integration into the platform years earlier.
- Other applications within the DoD include:
  - Simultaneous test programs (aircraft flight demonstrations and contract competitions).
  - Development of multiple and simultaneous weapons systems.

Portfolio Option (Basket of Options to Execute)

A portfolio of options provides the holder a variety or basket of possible option paths to execute. Some of these options may be too expensive, consistently be dominated by other options, take too long to execute, or simply be nonviable options. Determining the optimal portfolio of warfighter capabilities to develop and field within budgetary and time constraints is key to solving and modeling a portfolio optimization problem.

- Determine the optimal portfolios that provide the maximum capability, flexibility, and cost effectiveness with minimal risks given budget, schedule, wartime, and other scenarios. For instance, if Congress authorizes additional funding or cuts existing funding to certain programs, which capabilities or features should be added or cut?
- Helps to model and determine how much flexibility in design options should be incorporated into an MAS/FASO ship. Investing too little in flexibility will result in excessive modernization costs and increased downtime of the ship or its early retirement before the end of the design service life. Investing too much will create excess flexibility that will not be used, and create a higher up-front cost to obtain these flexibility options.
- Allows for different flexible pathways: Mutually Exclusive (C1 or C2 but not both), Platform Technology (C3 requires C2, but C2 can be standalone; expensive and worth less if considered by itself without accounting for flexibility downstream options it provides for the next phase), expansion options, abandonment options, and parallel development or simultaneous compound options.
- Other applications within the DoD include, but are not limited to:
  - Determining testing required in modular systems, mean-time-to-failure estimates, and replacement and redundancy requirements to maintain desired readiness and availability levels.
  - Maintaining capability and readiness at various levels.
  - Force mix options.
  - Capability selection and sourcing across a spectrum of vendors.

Sequential Compound Option (Proof of Concept, Milestones, and Stage-Gate Development)

The DoD has a requirement for an advanced technology to meet warfighter needs, but the technologies needed are in an early stage of maturity, and it is highly risky whether the technology will be available or work. There are limited vendors/activities capable of undertaking the development, so the program office may mitigate downside risks to the program through a phased approach to the acquisition. For instance, in the first phase, the vendor develops the underlying technology and presents the results to the PEO with a preliminary design. At the end of this phase, the government can either choose to continue through development of a prototype system or harvest the Science and Technology work for later use and abandon the effort. On delivery of a working prototype, the government will conduct tests for performance, evaluate total life-cycle cost, and decide whether to continue to full-scale system development or to abandon the effort, salvaging the knowledge from the prototyping effort for later use.

As an example, an acquisition program manager recognizes that multiple approaches to the problem are possible and decides to pursue a course of parallel development in which a variety of vendors and government labs undertake work to propose a technology solution, which creates a Multiple Activity or Multiple Vendor development of a system or technology. At option points (generally one to two years after contract award), the various solutions will be evaluated for performance, technical merit, and cost, and the universe of participants reduced through a down-select process. After two (or pick a number) rounds, the two most promising approaches are selected for advanced development and prototyping. From those, the best (evaluated in terms of performance, risk, and cost) will be selected for final development and fielding.

The U.S. Navy is currently pursuing the applications of new 3D scanning technology on board a ship to streamline the planning process for depot-level repair work. If the technology works after any technical problems have been ironed out, the scope can be expanded to implement online collaborative tools (requires additional investment) to implement additional process efficiencies for the management of depot-level ship repairs. Expansions across the population of Naval Shipyards will extend the savings/return on investment.

In pursuing Open Architecture (OA) over multiple stages, a proof of concept stage is performed first, and then several small-scale implementations and a final larger-scale implementation are executed. For instance, try OA modular development on a shore-based test system to see if it works before fielding on all units of that Class in the fleet once all the bugs are worked out and only if the proof of concept results are encouraging, thereby reducing the risk while at the same time obtaining the additional upside potential of going to OA (lower downtime, reduced cycle-time, reduced cost, interchangeable parts, at-sea repairs, multiple vendor parts for one system instead of relying only on a single vendor for the entire system, etc.). Successful implementation of a component or technology in one ship Class also provides the opportunity in an OA environment to expand to integrate the capability/technology into other open architected systems for other ship Classes.

A PM in charge of a large spiral development may need to determine the value of various items to release in each spiral. For example, the USAF logistics modernization program (called the Enterprise Resource Planning [ERP] System) has a goal to replace 250 separate legacy systems. A single release would likely be a huge failure. Developing various sequential strategies would show how to capture the most savings during each spiral release of the ERP system while minimizing risks as the system matures. The Army is also adopting the spiral development process for its logistics modernization program. Other examples of spiral development include the U.S. Air Force Air Theater Battle Management System and the Army's Future Combat Systems program, a system of systems development.

- Other applications within the DoD include, but are not limited to:
  - Stage-gate implementation of high-risk project development, prototyping, low-rate initial production (LRIP), technical feasibility tests, and technology demonstration competitions.
  - Government contracts with multiple stages with the option to abandon at any time and valuing Termination for Convenience (T-for-C), and built-in flexibility to execute different courses of action at specific stages of development.
  - P3I, Milestones, R&D, and Phased Options.
  - Platform technology development.

Expansion Option (Platform Technology With Spinoff Capabilities)

The C-17 Globemaster III is a long-range cargo/transport aircraft operated by the USAF since 1993. Full-scale development of the C-17 got underway in 1986, but technical problems and funding shortfalls delayed the program. Despite those difficulties, the C-17 retained broad support from Congress. In April 1990, Defense Secretary Cheney reduced the projected buy from 210 to 120 planes, exercising a contraction option. By the mid-1990s, the program's difficulties had been largely resolved. In 1996, the DoD approved plans for more C-17s and planned to end the production at 180 aircraft in FY2007. Congress then approved another $2 billion for 10 additional C-17 aircraft in FY2008. Expansion options put in place would allow the smooth addition of aircraft as needed, including foreign military sales. Other applications within the DoD include Platform Technologies, Acquisitions, ACTD Follow-on, Foreign Military Sales (FMS), Reusability and Scalability Options, and so forth.

Abandonment Option (Salvage and Walk Away)

A DoD research and development organization in conjunction with a military contractor decides to enter into a joint-testing agreement to test a satellite-based voice recognition intelligence gathering hardware-software product combination currently in its infancy stage of development that, if successful, could potentially be very useful in the fight against terrorism. The DoD can hedge its risks (i.e., the risk is the potential that the hardware-software combination will not work as required) and invest a small sum to buy the right of first refusal for a future investment, for some prespecified amount that is agreed upon now. This way, the U.S. Navy gets to participate in the technology if it is successful, but yet risks only a little if unsuccessful. In deciding whether to purchase the intelligence gathering equipment, a military analyst values the potential to abandon and sell off or divest the assets of the company in the future should there be no further use of the technology or if a newer and much more potent technology arrives on the market. The ability to do so will, in fact, reduce the risk on what the military has to spend on the technology and allows it to recoup some of its potential losses. Other applications within the DoD include Exit and Salvage (cutting losses), Stop before executing the next phase, Termination for Convenience (T-for-C), and so forth.

Contraction Option (Partnerships and Cost/Risk Reduction Strategy)

A contraction option allows two parties to create a joint venture or partnership (e.g., DoD and military vendor partnership) whereby the DoD agrees to purchase certain quantities of a product while holding partial intellectual property rights to the new development. Risks of failure are shared between the two parties, and no single party will bear all the risks (the DoD hedges its downside risks of the product failing, and the vendor hedges its risks of the DoD being not interested in its product). Other applications within the DoD include Outsourcing, Alliances, Contractors, Joint Inter-Service Venture, Foreign Partnerships, and so forth.

FASO/MAS at PEO-Ships: Flexibility Options for Guided Missile Destroyers

The Arleigh Burke class of Guided Missile Destroyers (DDG) is the U.S. Navy's first class of destroyer built around the Aegis Combat System and the SPY-1D multi-function passive electronically scanned array radar. The class is named for Admiral Arleigh Burke, the most famous American destroyer officer of World War II, and later Chief of Naval Operations. The class leader, USS *Arleigh Burke*, was commissioned during Admiral Burke's lifetime (Navy Programs, 2013).

The DDG class ships were designed as multi-mission destroyers to fit the Anti-Aircraft Warfare (AAW) role with their powerful Aegis radar and surface-to-air missiles; Anti-Submarine Warfare (ASW) with their towed sonar array, bow sonar, anti-submarine rockets, and ASW helicopter; Anti-Surface Warfare (ASUW) with their Harpoon missile launcher; and strategic land strike role with their Tomahawk missiles. With upgrades to AN/SPY-1 phased radar systems and their associated missile payloads as part of the Aegis Ballistic Missile Defense System, members of this class have also begun to demonstrate some promise as mobile anti-ballistic missile and anti-satellite weaponry platforms. Some versions of the class no longer have the towed sonar or Harpoon missile launcher (Navy Programs, 2013).

The DDG 51 class destroyers have been designed to support carrier strike groups, surface action groups, amphibious groups, and replenishment groups. They perform primarily AAW with secondary land attack, ASW, and ASUW capabilities. The Mk 41 vertical launch system has expanded the role of the destroyers in strike warfare, as well as their overall performance. The U.S. Navy will use the DDG 51 Flight III Destroyer equipped with the Aegis Modernization program and AMDR to provide joint battlespace threat awareness and defense capability to counter current and future threats in support of joint forces ashore and afloat. The following provides two high-level examples of identifying and framing strategic flexibility options in the DDG51 and DDG1000 environments. These are only notional examples with rough order magnitude values to illustrate the options framing approach.

Power Plant Options

This real options example illustrates the implications of the standard LM2500 GE Marine Gas Turbines for DDG51 FLT III ships versus the Rolls-Royce MT30 Marine Gas Turbine Engines for the Zumwalt DDG 1000, where the latter can satisfy large power requirements in warships. The LM2500 provides 105,000 shaft hp for a four-engine plant. In comparison, the MT30 can generate upwards of 35.4MW, and its auxiliary RR4500 Rolls-Royce turbine generators can produce an added 3.8MW, and each DDG1000 carries two MT30s and two RR4500s. This means that the combined energy output from the Zumwalt can fulfil the electricity demands in a small- to medium-size city. Manufacturer specifications indicate that the LM2500 has an associated Cost/kW of energy of $0.34 and the MT30 Cost/kW is $0.37. In addition, the MT30 prevents warships from running off balance when an engine cannot be restarted until it has cooled down, as is the case in the LM2500.

Figure 1 illustrates a real options strategy tree with four mutually exclusive paths. Additional strategies and pathways can similarly be created, but these initial strategies are sufficient to illustrate the options framing approach. Path 1 shows the As-Is strategy where no additional higher capacity power plant is used, that is, only two standard LM2500 units are deployed, maintain zero design margins for growth, and only what is required for the current ship configuration is designed and built. Medium and large upgrades will require major ship alterations, with high cost and delayed schedule. Path 2 implements the two required LM2500 units with additional and sufficient growth margins for one MT30 power plant but currently only with a smaller power plant incorporated into the design. Sufficient area or modularity is available where parts of the machinery can be removed and replaced with the higher energy production unit if needed. Upfront cost is reduced, and future cost and schedule delays are also reduced. Path 3 is to have two prebuilt MT30s and RR4500s initially. While providing the fastest implementation pathway, the cost is higher in the beginning, but total cost is lower if indeed higher energy weapons will be implemented. Path 4 is an option to switch whereby one LM2500 is built with one MT30 unit. Depending on conditions, either the LM2500 or MT30 will be used (switched between units). When higher-powered future weapons are required such as electromagnetic railguns (E.M. Rail Guns) or high-intensity lasers (H. I. Lasers) as well as other similarly futuristic weapons and systems, the MT30 can be turned on.

Having a warship flexibility with two LM2500s (As-Is base case), allows the Navy a savings of $31.76 million by deferring the option of the other two additional LM2500s. Therefore, having a flexible ship, the Navy can invest later in one LM2500 and attach another MT30 (preventing any engine off-balance effects when the engines cannot be restarted due to excessive heat), can save $34.58 million. The usage of options to defer/invest that combine gas turbine specifications allows the Navy to prevent high sunk costs, properly adjusting the true kW requirements, and allows different combinations of propulsion and energy plants. This analysis can be further extended into any direction as needed based on ship designs and Navy requirements.

The true competitor for MT-30 is LM2500+, as installed in LHD8, LHA6, and the trimaran/even-hull LCS. As an integrated electric plant, the analysis can also factor in ancillary generators (diesel or gas turbine) in addition to the main gas turbines to get an idea of total load capacity since loads can be shifted between propulsion and other uses. In contrast, our legacy plants can’t do that, although there is a move afoot to install auxiliary electric motors onto the reduction gears of the DDGs so they can slow-speed steam on electric power with mains offline.

Vertical Launch System

Another concern of the DoD is the large capital investments required in Vertical Launch Systems (VLS) in U.S. Navy ships. VLS need to be developed and integrated per Navy requirements, which are constrained by rapid technological change and high uncertainties in costs. The usage of strategic real options aims to assess whether the Navy can *keep the option open* to defer the large investments to help avoid high sunk costs and quick technological obsolescence, or whether the Navy should pre-invest in a new VLS. Consequently, flexibility and uncertainty create the right environment to model VLS using a real option framework. According to DGG 51 (Flight II and Flight III) specifications, the estimated cost of a single VLS is approximately $228 million. The most expensive subarea is the MK41 subsystem (DDG 51 contains two MK41s). This current example is developed based on the assumptions of a rapid technological obsolescence, high integration costs, time delays, and reduced capability, which can jeopardize investments in VLS.

In addition, using a real options framework to possibly defer the implementation of MK41 would allow ship designers and engineers to incorporate modernization and upgrade margins in the VLS within the ship design early on, and to defer the exact configuration of the VLS until a future date when uncertainties on capability requirements (i.e., integration, upgrades, changes, new technology, new requirements, updated military warfighter needs) are resolved over the passage of time, action, and events. Also, we can evaluate the option to invest in the second or third MK41 as the situational needs arise. Figure 2 shows the two simple option paths, in which the first path indicates immediate execution where two MK41s are implemented immediately, not knowing if both are actually needed, as opposed to the second strategic path where the VLS is designed such that either two MK41s can be implemented or only one. Therefore, one MK41 can be first inserted and the second added on later only when required, where the VLS has design growth margins to adapt to slightly different technological configurations. The question, of course, is which strategic pathway makes most sense, as computed using strategic real options value.

When the flexibility value is added into the mix, the expected total cost is reduced from $110.10 million to $98.51 million. Finally, wartime scenarios can be incorporated into the analysis whereby if there is a higher probability of conflict where the VLS is required, the value to keep open the option to defer is reduced, and the Navy is better off executing the option immediately and having the required VLS in place.

The project with flexibility is $118.22 million (flexible VLS warship open to integrate another MK41 in the future as and when needed) against $228.34 million (base case DDG 51 with no flexibility options, where the VLS is already built in). The Navy can save or delay the usage of $110.10 million in cost by holding on to the option of deferring the second MK41. In addition, in the near future, the cost to implement the second MK41 can be reduced due to a flatter learning curve, economies of scale, and the specific technology becoming more readily available, less complex, and easier to implement, or can be more expensive because the technology experiences new updates, higher performance, and greater efficiency. If cost volatility is the main variable for the Navy, we contrast differing the second MK41 against the base case. It means that we compare the VLS system with no flexibility ($228.32 million) against the cost changes in the second MK41 (assuming Navy engineers develop a plug-and-play structure to integrate the next MK41 quickly). This assumption can be relaxed using cost and schedule modeling and Monte Carlo simulation methods. In terms of the options valuation, the option to defer for the Navy follows cost comparisons. It other words, it reduces the cost exposure for the second MK41 from $110.10 million to an expected value of $69.89 million. In addition, decision-makers observe in the options strategy tree and decision tree where they can keep the option to defer open and under what conditions the Navy should execute and invest in the second MK41. One likely extension is where the decision-maker can introduce probabilities or expectations of Navy actions (new missions and new requirements) or events (wartime, peacetime). This affects the flexibility of the second MK41 by constraining the option’s flexibility to defer. For instance, if the Navy has strong expectations of requiring the second MK1 (wartime probability is higher than 30%), it reduces the value of the option to defer and accelerates the availability and execution of the second MK41 option earlier. In peacetime, the Navy has more flexibility in terms of how it implements or assesses its real options to wait and defer. Additional MK57 Peripheral VLS in use on DDG1000 or the MK48 family of self-defense VLS for ESSM can of course be additional considered.

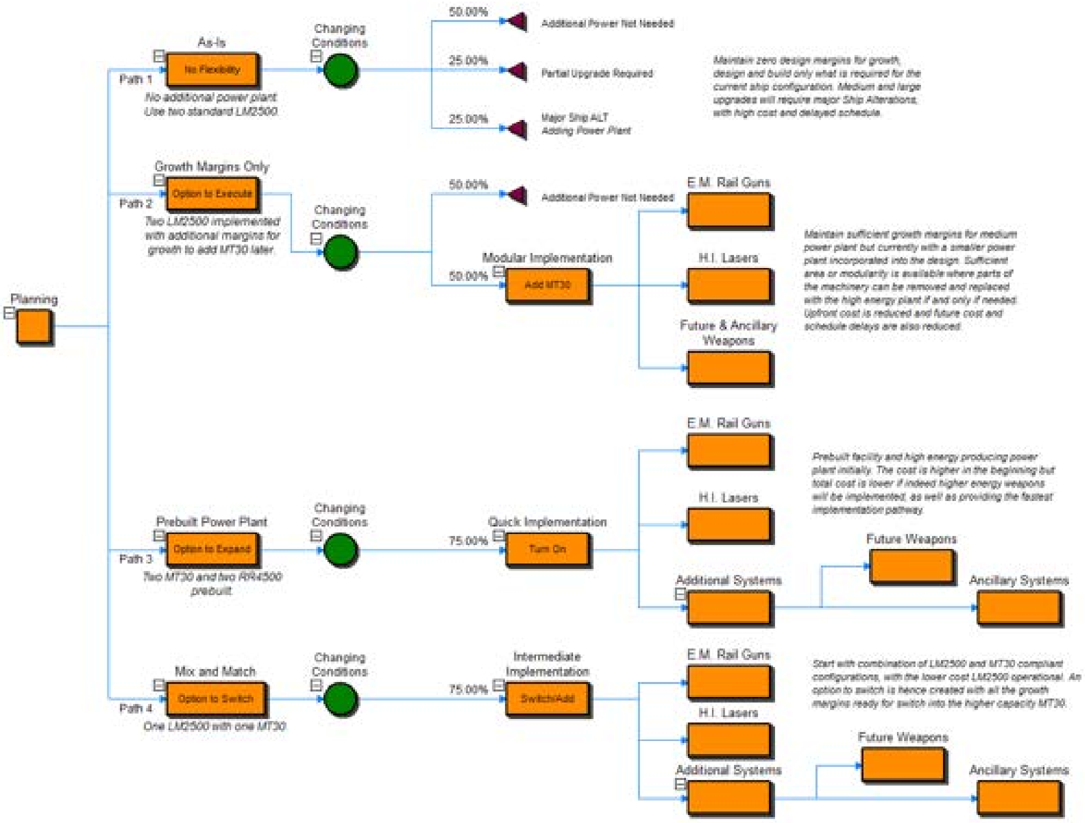


Figure 1. Options Framing on Power Generation

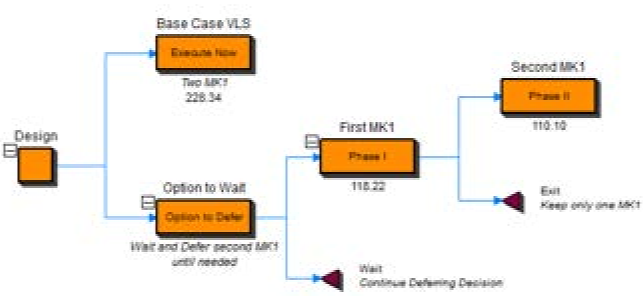


Figure 2. Options Framing on Vertical Launch Systems

The strategic real options analysis is solved employing various methodologies, including the use of binomial lattices with a market-replicating portfolios approach, and backed up using a modified closed-form sequential compound option model. The value of a compound option is based on the value of another option. That is, the underlying variable for the compound option is another option, and the compound option can be either sequential in nature or simultaneous. Solving such a model requires programming capabilities. This subsection is meant as a quick peek into the math underlying a very basic closed-form compound option. This section is only a preview of the detailed modeling techniques used in the current analysis and should not be assumed to be the final word. For instance, we first start by solving for the critical value of I, an iterative component in the model, using (Mun, 2016):

$$X_2 = Ie^{-q(T_2-t_1)}\Phi\left(\frac{\ln(I/X_1)+(r-q+\sigma^2/2)(T_2-t_1)}{\sigma\sqrt{(T_2-t_1)}}\right) - X_1 e^{-r(T_2-t_1)}\Phi\left(\frac{\ln(I/X_1)+(r-q-\sigma^2/2)(T_2-t_1)}{\sigma\sqrt{(T_2-t_1)}}\right)$$

Then, solve recursively for the value I above and input it into the model:

$$Compound\ Option = Se^{-qT_2}\Omega\left[\frac{\ln(S/X_1)+(r-q+\sigma^2/2)T_2}{\sigma\sqrt{T_2}};\ \frac{\ln(S/I)+(r-q+\sigma^2/2)t_1}{\sigma\sqrt{t_1}};\ \sqrt{t_1/T_2}\right]$$
$$-X_1 e^{-rT_2}\Omega\left[\frac{\ln(S/X_1)+(r-q+\sigma^2/2)T_2}{\sigma\sqrt{T_2}}-\sigma\sqrt{T_2};\ \frac{\ln(S/I)+(r-q+\sigma^2/2)t_1}{\sigma\sqrt{t_1}}-\sigma\sqrt{t_1};\ \sqrt{t_1/T_2}\right]$$
$$-X_2 e^{-rt_1}\Phi\left[\frac{\ln(S/I)+(r-q+\sigma^2/2)t_1}{\sigma\sqrt{t_1}}-\sigma\sqrt{t_1}\right]$$

The model is then applied to a sequential problem where future phase options depend on previous phase options (e.g., Phase II depends on Phase I's successful implementation).

Definitions of Variables

S	present value of future cash flows ($)	r	risk-free rate (%)
σ	volatility (%)	Φ	cumulative standard-normal
q	continuous dividend payout (%)	I	critical value solved recursively
Ω	cumulative bivariate-normal	X_1	strike for the underlying ($)
X_2	strike for the option on the option ($)	t_1	expiration date for the option on the option
T_2	expiration date for the underlying option		

The preceding closed-form differential equation models are then verified using the risk-neutral market-replicating portfolio approach assuming a sequential compound option. In solving the market-replicating approach, we use the following functional forms (Mun, 2016):

- Hedge ratio (h):
  $$h_{i-1} = \frac{C_{up} - C_{down}}{S_{up} - S_{down}}$$
- Debt load *(D)*
  $$D_{i-1} = S_i(h_{i-1}) - C_i$$
- Call value *(C)* at node i:
  $$C_i = S_i(h_i) - D_i e^{-rf(\delta t)}$$
- Risk-adjusted probability (q):
  $q_i = \frac{S_{i-1} - S_{down}}{S_{up} - S_{down}}$ obtained assuming
  $$S_{i-1} = q_i S_{up} + (1-q_i)S_{down}$$

- This means that
$S_{i-1} = q_i S_{up} + S_{down} - q_i S_{down}$ and $q_i [S_{up} - S_{down}] = S_{i-1} - S_{down}$
so we get $q_i = \frac{S_{i-1} - S_{down}}{S_{up} - S_{down}}$

Additional methods using closed-form solutions, binomial and trinomial lattices, and simulation approaches as well as dynamic simulated decision trees that are used in computing the relevant option values of each strategic pathways as previously indicated. Fortunately, Navy analysts do not have to be experts in advanced mathematics to run these models, as they have all been preprogrammed in PEAT, as illustrated in Figure 15.

Cost Analysis and Data Gathering

Once the various FASO/MASO options are framed and modeled as shown in the previous step, the modeling process continues with additional data gathering activities. The following are some sample parameters of the Surface Warfare program under consideration, and we use the generic terms Option 1, Option 2, and so forth, for generalization purposes.

- For all models, we can assume a 15% discount rate, 35% tax rate, and a 10-year time horizon for the cost savings (all future savings past Year 10 after discounting will be assumed to be negligible). The discounting base year is 2017 (Year 0 and Capital Investment is required in 2017) whereas immediate savings and short-term benefits and maintenance savings start in Year 1 (2018). This means Year 10 is 2027. These rates are applied only to monetary values and can be changed to whatever appropriate values as required.
- The following table shows the remaining relevant information you will need to run your models. All monetary values are in thousands of dollars ($000). Remember to save your models and settings.

Capability Options	Savings Now	Short-Term Benefits	Maintenance Savings	Capital Cost	Fixed Cost	Operating Cost	OPNAV Value	Command Value	KVA Value
Option 1	$550	$30	$60	$400	$3	$2	8.1	1.2	8.1
Option 2	$650	$5	$10	$300	$3	$2	1.27	2.5	1.27
Option 3	$700	$35	$10	$350	$3	$2	5.02	7.5	5.02
Option 4	$1,000	$50	$20	$600	$3	$2	8.83	4.5	8.83
Option 5	$2,000	$100	$20	$1,000	$3	$2	9.88	9.7	9.88
Option 6	$1,000	$10	$20	$550	$3	$2	3.64	7.4	3.64
Option 7	$2,000	$100	$20	$750	$3	$2	5.27	4.5	5.27
Option 8	$850	$75	$20	$550	$3	$2	9.8	7.5	9.8
Option 9	$1,500	$125	$20	$750	$3	$2	5.68	7.5	5.68
Option 10	$1,000	$125	$20	$550	$3	$2	8.29	8.5	8.29

- “Savings Now” is the immediate monetary cost savings benefits obtained by implementing the new upgraded system (e.g., lower overhead requirements, reduced parts and labor requirements). This amount is applied in the first year of the cash flow stream only (Year 1 or 2016) as its effects are deemed as immediate.
- “Short-Term Benefits” is the savings per year for the first 5 years, stemming from reduction in staffing requirements, but these savings are deemed to be reabsorbed later on. Savings apply from 2016 to 2020.
- “Maintenance Savings” is the savings each year for all 10 years starting in 2016 where system maintenance cost is reduced and saved.
- “Capital Cost” is applied in Year 0 or 2015 as a one-time capital expenditure.
- Assume a “Fixed Direct Cost” and constant “Indirect Operating Cost” per year for all 10 years starting in 2016. The new equipment upgrades will require

some fixed overhead cost and operating expenses to maintain. The idea is these will be less than the total sum of benefits obtained by implementing the capability.

- "OPNAV" and "Command" are average values of multiple subject matter experts' estimates of the criticality (1–10, with 10 being the highest) of each capability. "KVA" is unit equivalence (this can be multiplied by any market price comparable such as $1 million per unit or used as-is in the optimization model). These will be used later in the optimization section.

Financial Modeling

The *Discounted Cash Flow* section shown in Figure 3 is at the heart of the analysis's input assumptions. Users would enter their input assumptions—such as starting and ending years of the analysis, the discount rate to use, and the marginal tax rate—and set up the project economics model (adding or deleting rows in each subcategory of the financial model). Additional time-series inputs are entered in the data grid as required, while some elements of this grid are intermediate computed values. The entire grid can be copied and pasted into another software application such as Microsoft Excel, Microsoft Word, or other third-party software applications, or can be viewed in its entirety as a full screen pop-up.

Users can also identify and create the various options, and compute the economic and financial results such as net present value (NPV), internal rate of return (IRR), modified internal rate of return (MIRR), profitability index (PI), return on investment (ROI), payback period (PP), and discounted payback (DPP). This section will also auto-generate various charts, cash flow ratios and models, intermediate calculations, and comparisons of the options within a portfolio view, as illustrated in the next few figures. As a side note, the term *Project* is used in PEAT's DCF module to represent a generic analysis option, where each project can be a different asset, project, acquisition, investment, research and development, or simply variations of the same investment (e.g., different financing methods when acquiring the same firm, different market conditions and outcomes, or different scenarios or implementation paths). Therefore, the more flexible terminology of *Project* is adopted instead.

Figure 4 illustrates the *Economic Results* of each project. This Level 3 subtab shows the results from the chosen project and returns the NPV, IRR, MIRR, PI, ROI, PP, and DPP. These computed results are based on the user's selection of the discounting convention, if there is a constant terminal growth rate, and the cash flow to use (e.g., net cash flow versus net income or operating cash flow). An *NPV Profile* table and chart are also provided, where different discount rates and their respective NPV results are shown and charted. Users can change the range of the discount rates to show/compute by entering the From/To percent, copy the results, and copy the profile chart, as well as use any of the chart icons to manipulate the chart's look and feel (e.g., change the chart's line/background color, chart type, chart view, or add/remove gridlines, show/hide labels, and show/hide legend). Users can also change the variable to display in the chart. For instance, users can change the chart from displaying the NPV profile to the time-series charts of net cash flows, taxable income, operating cash flows, cumulative final cash flows, present value of the final cash flows, and so forth.

The *Economic Results* subtabs are for each individual project, whereas *the Portfolio Analysis* tab (which is shown later as Figure 5) compares the economic results of all projects at once. The *Terminal Value Annualized Growth Rate* is applied to the last year's cash flow to account for a perpetual constant growth rate cash flow model, and these future cash flows, depending on which cash flow type chosen, are discounted back to the base year and added to the NPV to arrive at the perpetual valuation.

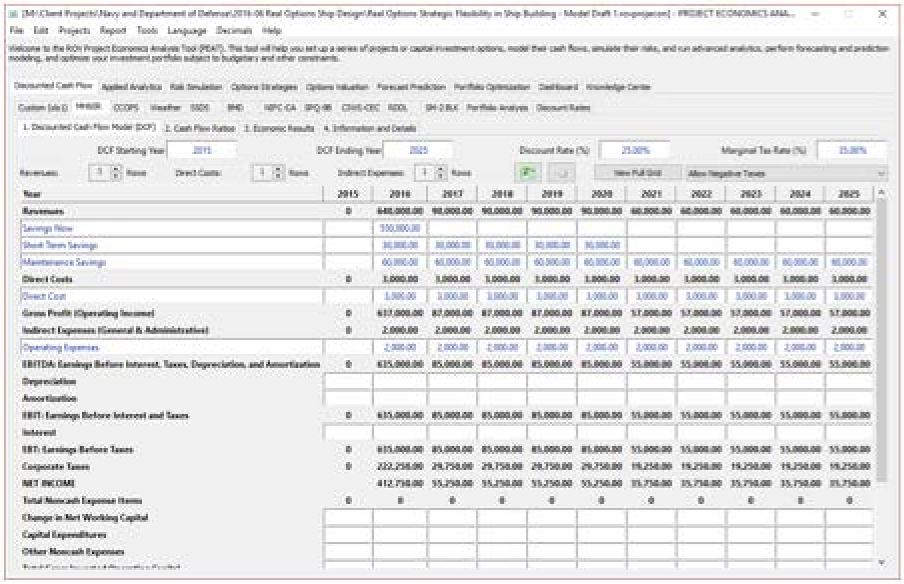


Figure 3. PEAT Discounted Cash Flow Module

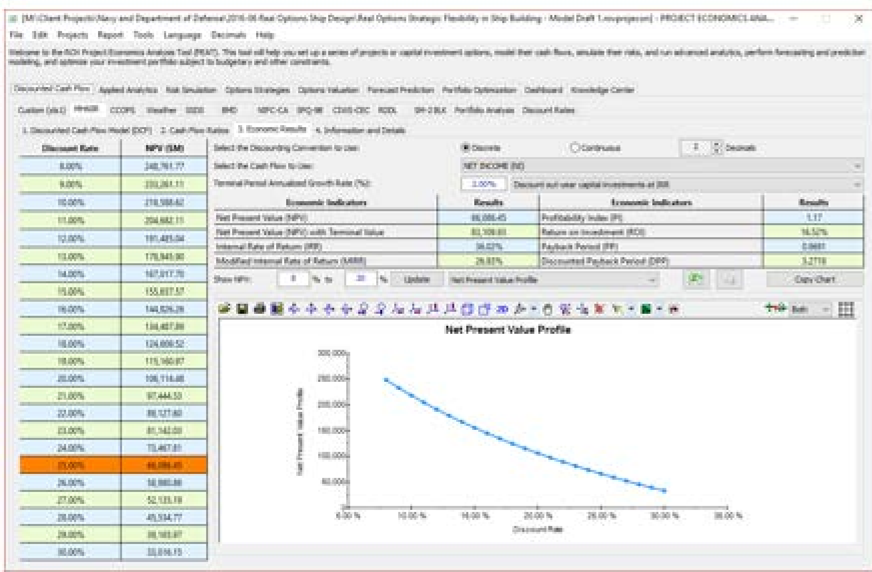


Figure 4. Economic Results

Static Portfolio Analysis and Comparisons of Multiple Projects

Figure 5 illustrates the *Portfolio Analysis* of multiple *Projects*. This Portfolio Analysis tab returns the computed economic and financial indicators such as NPV, IRR, MIRR, PI, ROI, PP, and DPP for all the projects combined into a portfolio view (these results can be stand-alone with no base case or computed as incremental values above and beyond the chosen base case). The *Economic Results* (Level 3) subtabs show the individual project's economic and financial indicators, whereas this Level 2 *Portfolio Analysis* view shows the results of all projects' indicators and compares them side by side. There are also two charts available for comparing these individual projects' results. The *Portfolio Analysis* tab is used to obtain a side-by-side comparison of all the main economic and financial indicators of all the projects at once. For instance, users can compare all the NPVs from each project in a single results grid. The bubble chart on the left provides a visual representation of up to three chosen variables at once (e.g., the y-axis shows the IRR, the x-axis represents the NPV, and the size of the bubble may represent the capital investment; in such a situation, one would prefer a smaller bubble that is in the top right quadrant of the chart). These charts have associated icons that can be used to modify their settings (chart type, color, legend, etc.).

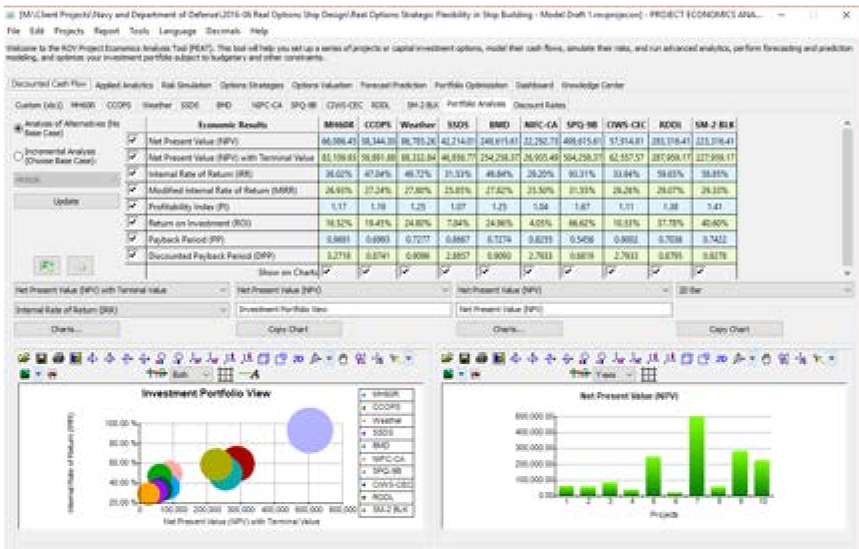


Figure 5. Static Portfolio Analysis

Tornado and Sensitivity Analytics

Figure 6 illustrates the *Applied Analytics* section, which allows users to run *Tornado Analysis* and *Scenario Analysis* on any one of the projects previously modeled—this analytics tab is on Level 1, which means it covers all of the various projects on Level 2. Users can, therefore, run tornado or scenario analyses on any one of the projects. Tornado analysis, as we already know, is a static sensitivity analysis of the selected model's output to each input assumption, performed one at a time, and ranked from most impactful to the least. Users start the analysis by first choosing the output variable to test from the droplist.

Users can change the default sensitivity settings of each input assumption to test and decide how many input variables to chart (large models with many inputs may generate unsightly and less useful charts, whereas showing just the top variables reveals more information through a more elegant chart). Users can also choose to run the input assumptions as unique inputs, group them as a line item (all individual inputs on a single line item are assumed to be one variable), or run as variable groups (e.g., all line items under Revenue will be assumed to be a single variable). Users will need to remember to click Update to run the analysis if they make any changes to any of the settings. The sensitivity results are also shown as a table grid at the bottom of the screen (e.g., the initial base value of the chosen output variable, the input assumption changes, and the resulting output variable's sensitivity results). The following summarizes the tornado analysis chart's main characteristics:

- Each horizontal bar indicates a unique input assumption that constitutes a precedent to the selected output variable.
- The x-axis represents the values of the selected output variable. The wider the bar chart, the greater the impact/swing the input assumption has on the output.
- A green bar on the right indicates that the input assumption has a positive effect on the selected output (conversely, a red bar on the right indicates a negative effect).
- Each of the precedent or input assumptions that directly affect the NPV with Terminal Value is tested ±10% by default (this setting can be changed); the top 10 variables are shown on the chart by default (this setting can be changed), with a 2-decimal precision setting; and each unique input is tested individually.
- The default sensitivity is globally ±10% of each input variable but each of these inputs can be individually modified in the data grid. Note that a larger percentage variation will test for nonlinear effects as well.
- The model's granularity can be set (e.g., Variable Groups look at an entire variable group such as all revenues or direct costs will be modified at once; Line Items change the entire row for multiple years at once; and Individual Unique Inputs look at modifying each input cell).

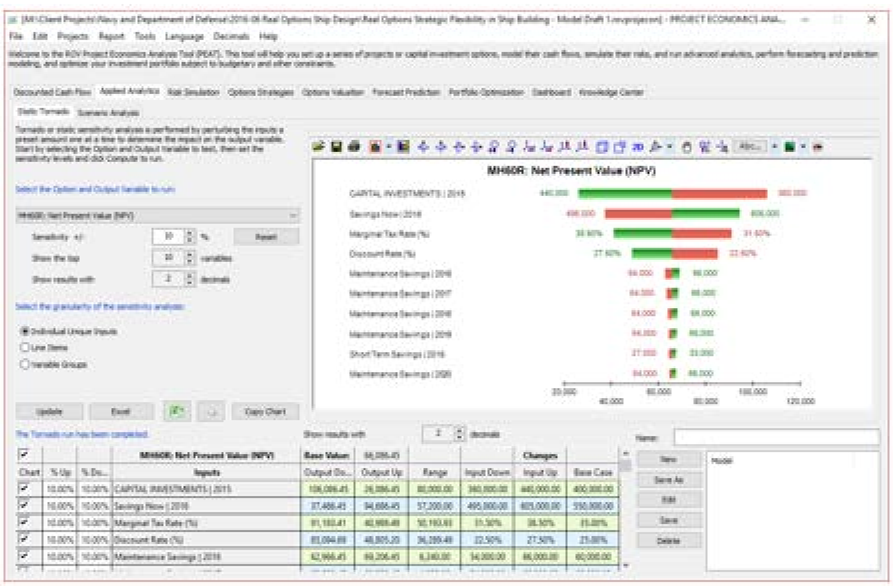


Figure 6. Applied Analytics: Tornado

Figure 7 illustrates the *Scenario Analysis* tab, where the scenario analysis can be easily performed through a two-step process: identify the model input settings and run the model to obtain scenario output tables. In the *Scenario Input Settings* subtab, users start by selecting the output variable they wish to test from the droplist. Then, based on the selection, the precedents of the output will be listed under two categories (*Line Item*, which will change all input assumptions in the entire line item in the model simultaneously, and *Single Item*, which will change individual input assumption items). Users select one or two checkboxes at a time and the inputs they wish to run scenarios on, and enter the plus/minus percentage and the number of steps between these two values to test. Users can also add color coding of sweetspots or hotspots in the scenario analysis (values falling within different ranges have unique colors). Users can create multiple scenarios and *Save As* each one (enter a name and model notes for each saved scenario).

Scenario analyses can sometimes be used as heat maps to identify the combinations of input parameter conditions whereby the calculated outputs will be above or below certain thresholds. A visual heat map can be created by adding color thresholds in the scenario results table. Figure 8 illustrates the *Scenario Output Tables* to run the saved *Scenario Analysis* models.

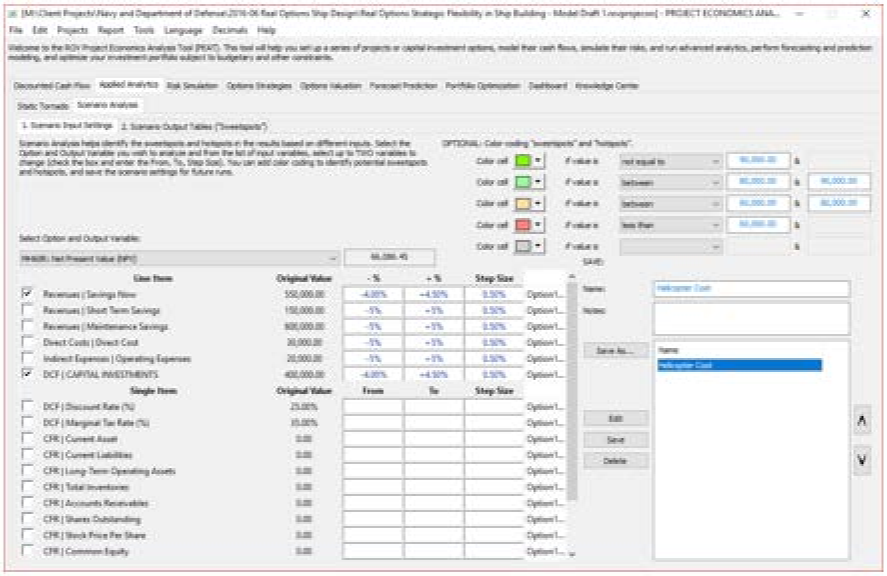


Figure 7. Applied Analytics: Scenario Analysis Input

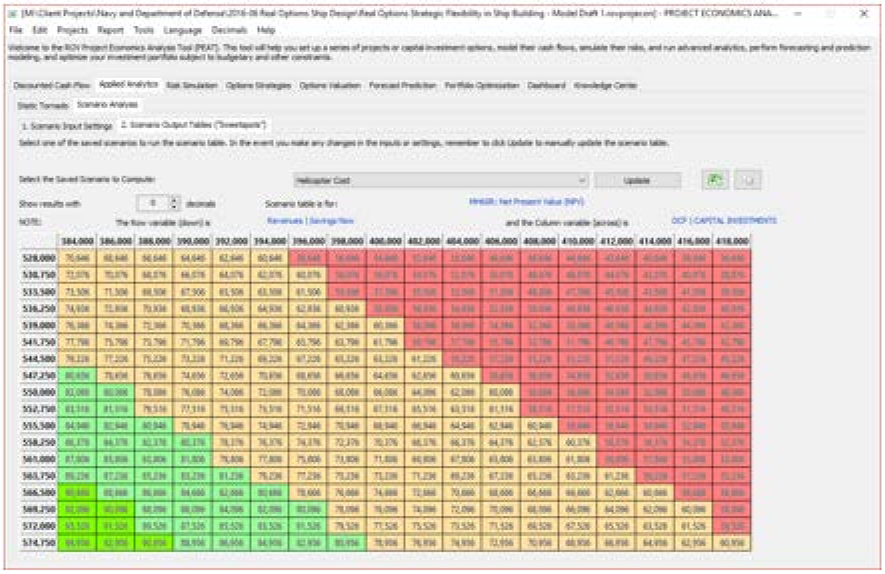


Figure 8. Applied Analytics: Scenario Tables

Monte Carlo Risk Simulation

Figure 9 illustrates the *Risk Simulation* section, where Monte Carlo risk simulations can be set up and run. Users can set up probability distribution assumptions on any combinations of inputs, run a risk simulation tens to hundreds of thousands of trials, and retrieve the simulated forecast outputs as charts, statistics, probabilities, and confidence intervals in order to develop comprehensive risk profiles of the projects.

Simulation Results, Confidence Intervals, and Probabilities

Figure 10 illustrates the Risk Simulation results. After the simulation completes its run, the utility will automatically take the user to the *Simulation Results* tab. The user selects the output variable to display using the droplist. The simulation forecast chart is shown on the left, while percentiles and simulation statistics are presented on the right.

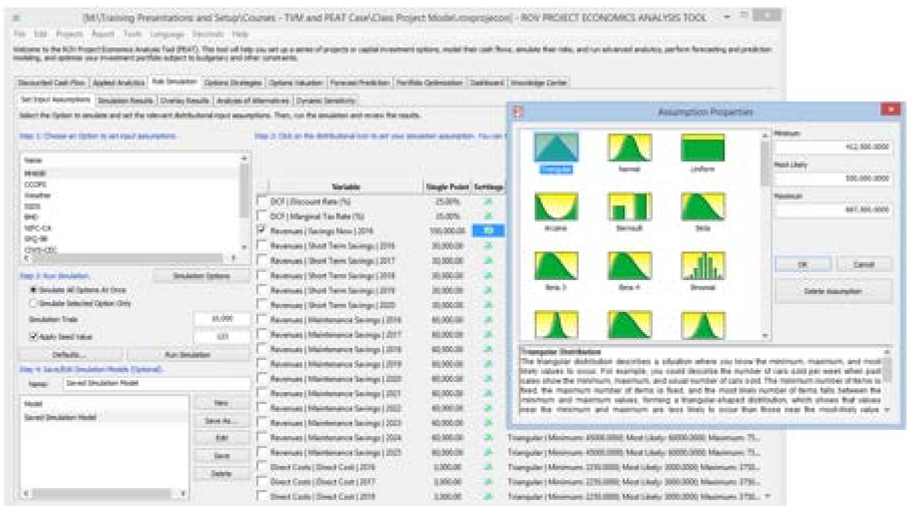


Figure 9. Risk Simulation Input Assumptions

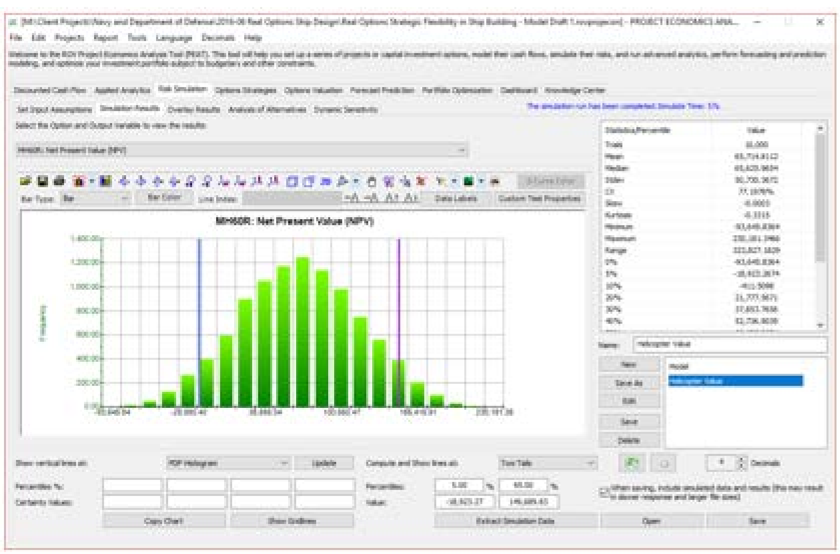


Figure 10. Risk Simulation Results

Probability Distribution Overlay Charts

Figure 11 illustrates the Overlay Results tab. Multiple simulation output variables can be compared at once using the overlay charts. Users simply check/uncheck the simulated outputs they wish to compare and select the chart type to show (e.g., S-Curves, CDF, PDF). Users can also add percentile or certainty lines by first selecting the output chart, entering the relevant values, and clicking the *Update* button. As usual, the generated charts are highly flexible in that users can modify them using the included chart icons (as well as whether to show or hide gridlines), and the chart can be copied into the Microsoft Windows clipboard for pasting into another software application. Typically, S-curves of CDF curves are used in overlay analysis when comparing the risk profile of multiple simulated forecast results.

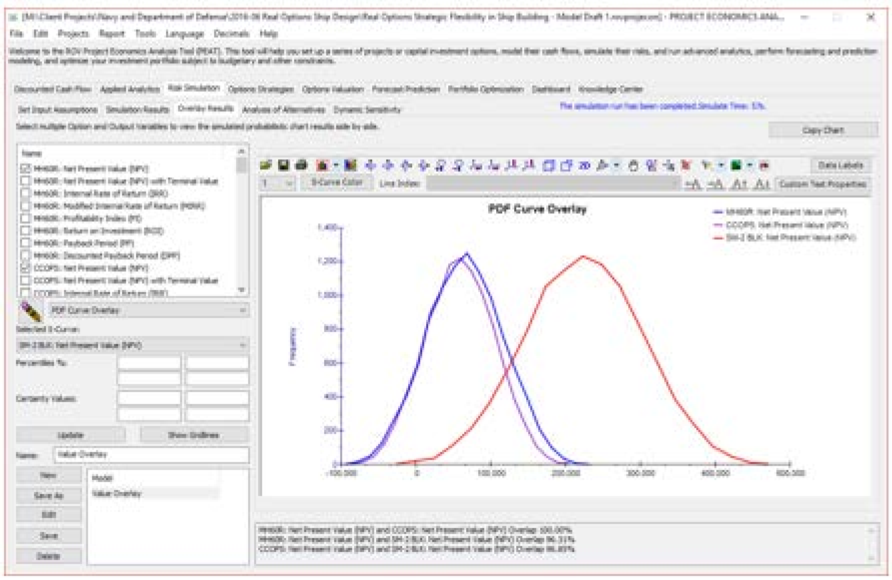


Figure 11. Simulated Overlay Results

Analysis of Alternatives

Figure 12 illustrates the *Analysis of Alternatives* subtab. Whereas the *Overlay Results* subtab shows the simulated results as charts (PDF/CDF), the *Analysis of Alternatives* subtab shows the results of the simulation statistics in a table format as well as a chart of the statistics such that one project can be compared against another. The default is to run an analysis of alternatives to compare one project versus another, but users can also choose the *Incremental Analysis* project (remembering to choose the desired economic metric to show, its precision in terms of decimals, the *Base Case* project to compare the results to, and the chart display type).

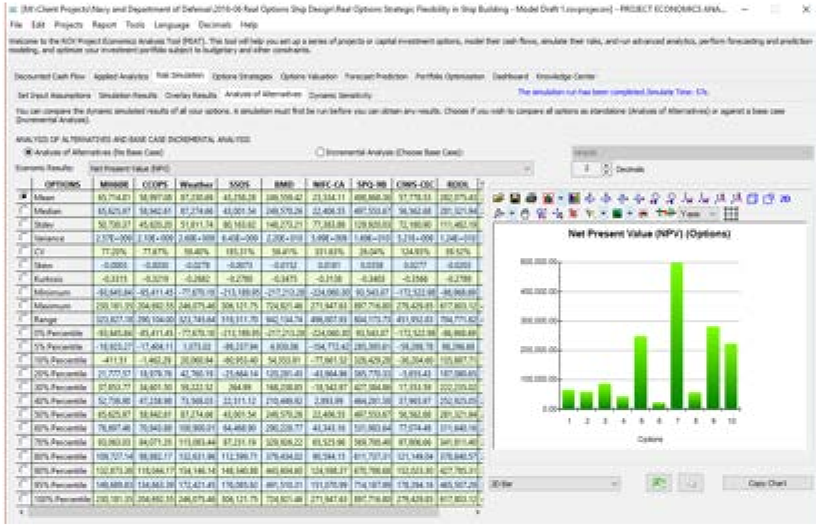


Figure 12. Simulated Analysis of Alternatives

Dynamic Sensitivity Analysis

Figure 13 illustrates the *Dynamic Sensitivity Analysis* computations. Tornado analysis and scenario analysis are both static calculations. Dynamic sensitivity, in contrast, is a dynamic analysis, which can only be performed after a simulation is run. Users start by selecting the desired project's economic output. Red bars on the *Rank Correlation* chart indicate negative correlations and green bars indicate positive correlations for the left chart. The correlations' absolute values are used to rank the variables from the highest relationship to the lowest, for all simulation input assumptions. The *Contribution to Variance* computations and chart indicate the percentage fluctuation in the output variable that can be statistically explained by the fluctuations in each of the input variables. As usual, these charts can be copied and pasted into another software application.

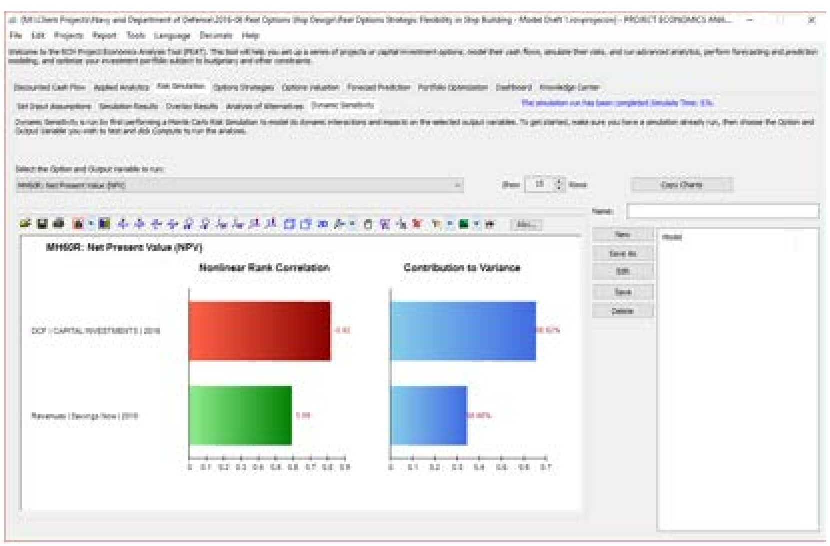


Figure 13. Simulated Dynamic Sensitivity Analysis

Strategic Real Options Valuation Modeling

Figure 14 illustrates the Options Strategies tab. Options Strategies is where users can draw their own custom strategic maps, and each map can have multiple strategic real options paths. This section allows users to draw and visualize these strategic pathways and does not perform any computations.

Real Options Valuation Modeling

Figure 15 illustrates the *Options Valuation* tab and the *Strategy View*. This section performs the calculations of real options valuation models. Users must understand the basic concepts of real options before proceeding. This *Options Valuation* tab internalizes the more sophisticated Real Options SLS. Instead of requiring more advanced knowledge of real options analysis and modeling, users can simply choose the real option types, and the required inputs will be displayed for entry. Users can compute and obtain the real options value quickly and efficiently, as well as run the subsequent tornado, sensitivity, and scenario analyses.

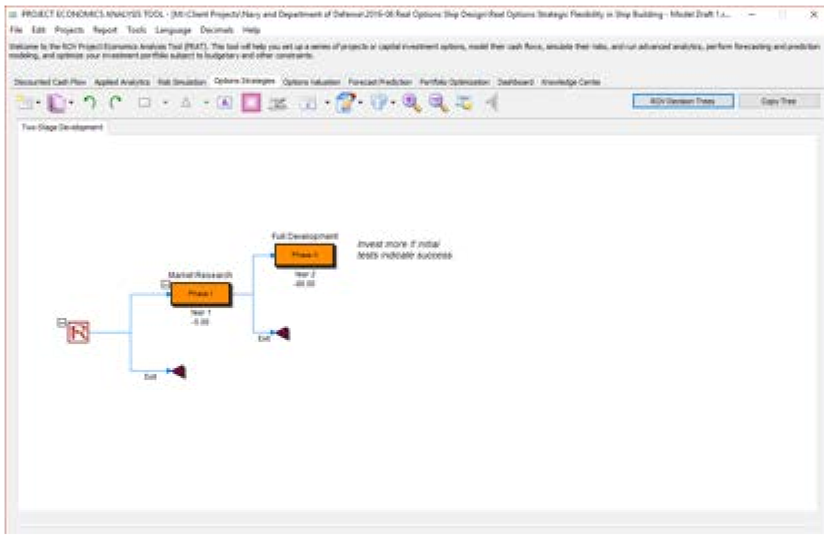


Figure 14. Options Strategies

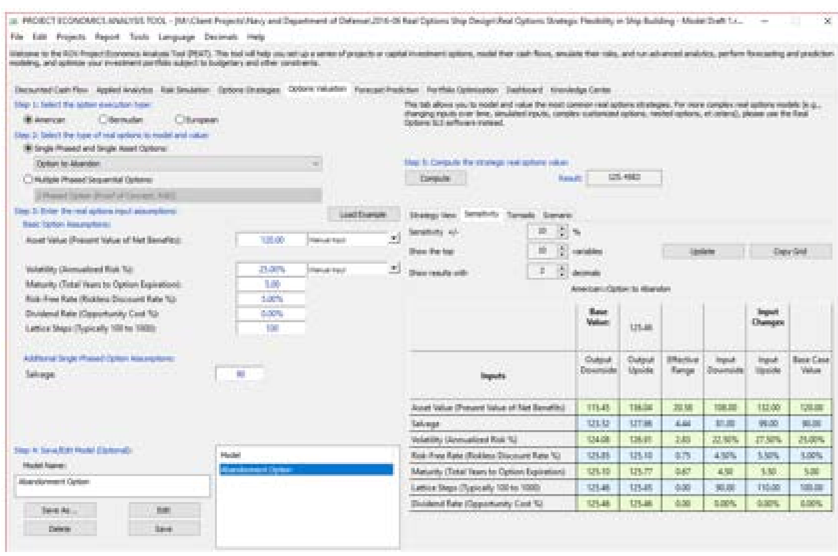


Figure 15. Options Valuation

Portfolio Optimization

Figure 16 illustrates the *Portfolio Optimization's Optimization Settings* subtab. In the Portfolio Optimization section, the individual projects can be modeled as a portfolio and optimized to determine the best combination of projects for the portfolio. In today's competitive global economy, companies are faced with many difficult decisions. These decisions include allocating financial resources, building or expanding facilities, managing inventories, and determining product-mix strategies. Such decisions might involve thousands or millions of potential alternatives. Considering and evaluating each of them would be impractical or even impossible. A model can provide valuable assistance in incorporating relevant variables when analyzing decisions and in finding the best solutions for making decisions. Models capture the most important features of a problem and present them in a form that is easy to interpret. Models often provide insights that intuition alone cannot. An optimization model has three major elements: decision variables, constraints, and an objective. In short, the optimization methodology finds the best combination or permutation of decision variables (e.g., which products to sell or which projects to execute) in every conceivable way such that the objective is maximized (e.g., revenues and net income) or minimized (e.g., risk and costs) while still satisfying the constraints (e.g., budget and resources).

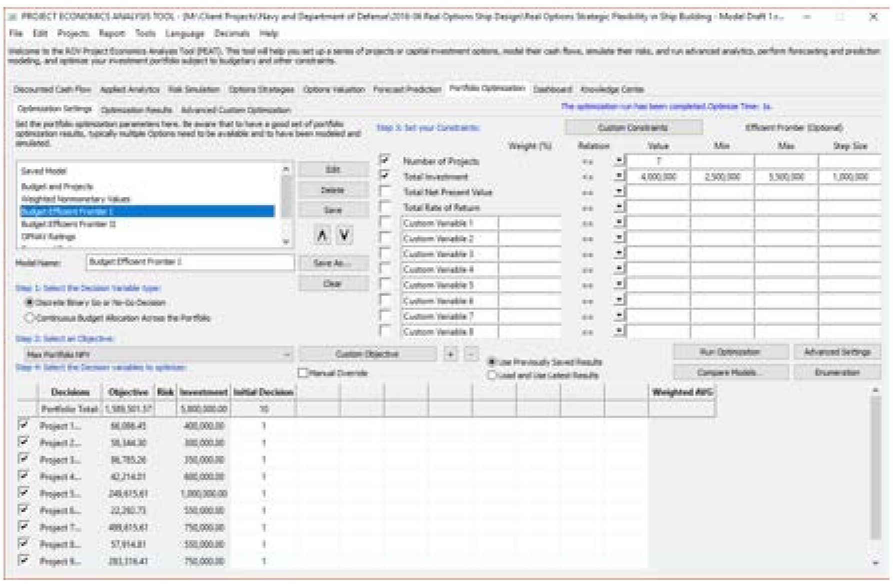


Figure 16. Portfolio Optimization Settings

The projects can be modeled as a portfolio and optimized to determine the best combination of projects for the portfolio in the *Optimization Settings* subtab. Users start by selecting the optimization method (Static or Dynamic Optimization). Then they select the decision variable type of *Discrete Binary* (choose which Project or Options to execute with a Go/No-Go Binary 1/0 decision) or *Continuous Budget Allocation* (returns % of budget to allocate to each option or project as long as the total portfolio is 100%); select the *Objective* (Max NPV, Min Risk, etc.); set up any *Constraints* (e.g., budget restrictions, number of projects restrictions, or create customized restrictions); select the options or projects to optimize/allocate/choose (default selection is *all options*); and when completed, click *Run Optimization.*

Figure 17 illustrates the *Optimization Results* tab, which returns the results from the portfolio optimization analysis. The main results are provided in the data grid, showing the final Objective Function results, final *Optimized Constraints*, and the allocation, selection, or optimization across all individual options or projects within this optimized portfolio. The top left portion of the screen shows the textual details and results of the optimization algorithms applied, and the chart illustrates the final objective function. The chart will only show a single point for regular optimizations, whereas it will return an investment efficient frontier curve if the optional *Efficient Frontier* settings are set (min, max, step size) in the tab.

Figures 17 and 18 are critical results for decision-makers as they allow flexibility in designing their own portfolio of options. For instance, Figure 17 shows an efficient frontier of portfolios, where each of the points along the curve are optimized portfolios subject to a certain set of constraints. In this example, the constraints were the number of options that can be selected in a ship and the total cost of obtaining these options are subject to a budget constraint. The colored columns on the right in Figure 17 show the various combinations of budget limits and maximum number of options allowed. For instance, if a program office in the Navy only allocates $2.5 million (see the Frontier Variable located on the second row) and no more than four options per ship, then only options 3, 7, 9, and 10 are feasible, and this portfolio combination would generate the highest bang for the buck while simultaneously satisfying the budgetary and number of options constraints. If the constraints were relaxed to, say, five options and $3.5 million budget, then option 5 is added to the mix. Finally, at $4.5 million and no more than seven options per ship, options 1 and 2 should be added to the mix. Interestingly, even with a higher budget of $5.5 million, the same portfolio of seven options is selected. In fact, the Optimized Constraint 2 shows that only $4.1 million is used. Therefore, as a decision-making tool for the budget-setting officials, the maximum budget that should be set for this portfolio of options should be $4.1 million. Similarly, the decision-maker can move backwards, where say, if the original budget of $4.5 million was slashed by the U.S. Congress to $3.5 million, then the options that should be eliminated would be options 1 and 2.

While Figure 17 shows the efficient frontier where the constraints such as number of options allowed and budget were varied to determine the efficient portfolio selection, Figure 18 shows multiple portfolios with different objectives. For instance, the five models shown were to maximize the financial bang for the buck (minimizing cost and maximizing value while simultaneously minimizing risk), maximizing OPNAV value, maximizing KVA value, maximizing Command value, and maximizing a Weighted Average of all objectives. This capability is important because depending on who is doing the analysis, their objectives and decisions will differ based on different perspectives. Using a multiple criteria optimization approach allows us to see the scoring from all perspectives. Options with the highest count (e.g., 5) would receive the highest priority in the final portfolio, because it satisfies all stakeholders' perspectives, and would hence be considered first, followed by options with counts of 4, 3, 2, and 1.

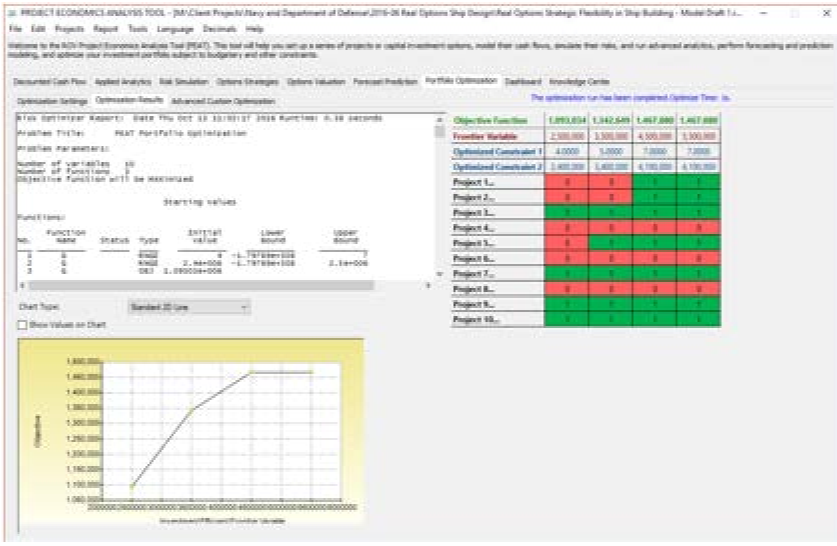


Figure 17. Portfolio Optimization Results

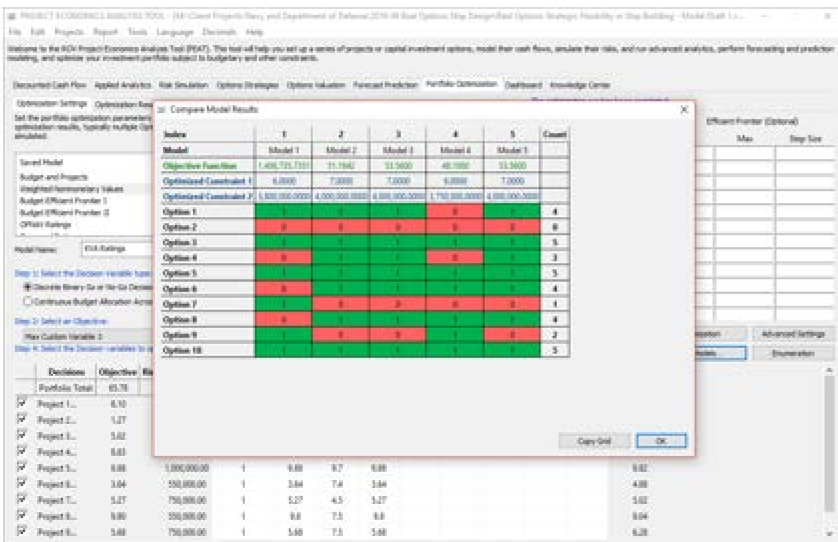


Figure 18. Multicriteria Portfolio Optimization Results

Analytics

As a side note and for the purposes of being comprehensive and inclusive, we point out that multiple types of algorithms have been developed over the years to find the solutions of an optimization problem, from basic linear optimization using the simplex model and solving first partial differential equations. However, when more and more complex real-life problems are assumed, these basic methods tend to break down, and more advanced algorithms are required. In solving our efficient frontier problem, we utilized a combination of genetic algorithm, Lagrange multipliers, and taboo-based reduced gradient search methodologies.

Simplistically, the Lagrange multiplier solution assumes some nonlinear problem of:

$$\min or \max f(x)$$

$$s.t. \quad g_i(x) = b_i \ \forall\, i = 1, \dots, m$$

where the equality is oftentimes replaced by some inequality values indicating a ceiling or floor constraint.

From this functional form, we first derive the Lagrange multiplier v for all i values:

$$L(x, v) \triangleq f(x) + \sum_{i=1}^{m} v_i[b_i - g_i(x)]$$

$$s.t. constraints\ g_i(x) = b_1, \dots, g_m(x) = b_m$$

The solution $(x*, v*)$ is a set of points along the Lagrange function $L(x, v)$ if it satisfies the condition:

$$\sum_i \nabla g_i(x^*)v^* = f(x^*)\ which\ requires\ \sum_i \frac{\partial g_i}{\partial x_j} v_i = \frac{\partial f}{\partial x_j} \forall j\ and\ \ g_i(x^*) = b_i$$

This approach is simple and elegant but limited to linear and quasi-linear as well as some simple nonlinear functional forms of $f(x)$. In order to be able to extend the functional form to generalized nonlinear applications, we need to add additional conditions to the solution set and apply some search algorithms to cover a large (and oftentimes unlimited set of optimal allocations). One limitation is the requirement that the Kuhn-Tucker condition is satisfied where the nonlinear problems have a differentiable general form:

$$\min or \max f(x)$$

$$\begin{aligned} s.t. \quad & g_i(x) \geq\ b_i \ \forall\, i \in Feasible\ Set \\ & g_i(x) \leq\ b_i \ \forall\, i \in Feasible\ Set \\ & g_i(x) =\ b_i \ \forall\, i \in Feasible\ Set \end{aligned}$$

and the inequality constraints will need to be active at a local optimum or when the Lagrange variable is set to null:

$$v_i[b_i - g_i(x)] = 0$$

In addition, mathematical algorithms will have to be developed to perform both an ad-hoc and systematic search of the optimal solution set. Using an enumeration method will take even a supercomputer close to an infinite number of years to delineate all possible permutations. Therefore, search algorithms are typically used in generating an efficient frontier using optimization. One simple approach is the use of a reduced gradient search method. To summarize the approach, we assume

$$\nabla f(x) \cdot \Delta x$$

where the functional form $f(x)$ is the objective function and is divided into two parts, a basic (B) and non-basic portion (*N*) is multiplied by the change in vector direction x. Using a Taylor expansion, we obtain:

$$\nabla f(x) \cdot \Delta x = \nabla f(x)^B \cdot \Delta x^B + \nabla f(x)^N \cdot \Delta x^N$$

$$= \nabla f(x)^B \cdot (-B^{-1}N\Delta x^N) + \nabla f(x)^N \cdot \Delta x^N$$
$$= (\nabla f(x)^N - \nabla f(x)^B B^{-1}N)\Delta x^N$$

The reduced gradient with respect to the solution matrix B is:

$$r \triangleq (r^B, r^N)$$

where

$$r^B \triangleq 0$$
$$r^N \triangleq \nabla f(x)^N - \nabla f(x)^B B^{-1}N$$

Solving for this solution set is manually possible when the number of decision variables is small (typically less than four or five), but once the number of decision variables is large, as in all real-life situations, the manual solution is intractable and computer search algorithms have to be employed.

Conclusions and Recommendations

First, it is vital to understand that real options analysis is *not* a simple set of equations or models. It is an *entire decision-making process* that enhances the traditional decision analysis approaches. It takes what has been tried-and-true financial analytics and evolves it to the next step by pushing the envelope of analytical techniques. In addition, it is vital to understand that 50% of the value in real options analysis is simply thinking about it. Another 25% of the value comes from the number crunching activities, while the final 25% comes from the results interpretation and explanation to management. Several issues should be considered when attempting to implement real options analysis:

- **Tools**—The correct tools are important. These tools must be more comprehensive than initially required because analysts will grow into them over time. Do not be restrictive in choosing the relevant tools. Always provide room for expansion. Advanced tools will relieve the analyst of detailed model-building and let him or her focus instead on 75% of the value—thinking about the problem and interpreting the results.
- **Resources**—The best tools in the world are useless without the relevant human resources to back them up. Tools do not eliminate the analyst but enhance the analyst’s ability to effectively and efficiently execute the analysis. The right people with the right tools will go a long way. Because there are only a few true real options experts in the world who truly understand the theoretical underpinnings of the models as well the practical applications, care should be taken in choosing the correct team. A team of real options experts is vital in the success of the initiative. A company should consider building a team of in-house experts to implement real options analysis and to maintain the ability for continuity, training, and knowledge transfer over time. Knowledge and experience in the theories, implementation, training, and consulting are the core requirements of this team of individuals. This is why training is vital. For instance, the CRM/CQRM certification program provides analysts and managers the opportunity to immerse themselves into the theoretical and real-life applications of simulation, forecasting, optimization, and real options (for details, please see www.realoptionsvaluation.com).
- **Senior Decision-Maker Buy-in**—The analysis buy-in has to be top-down where senior management drives the real options analysis initiative. A bottom-up approach where a few inexperienced junior analysts try to impress the powers that be will fail miserably.

References

Doerry, N. H. (2012, August). Institutionalizing modular adaptable ship technologies. *Journal of Ship Production and Design, 30*(3), 126–141.

Ford, D., Housel, T., & Mun, J. (2011). *Ship maintenance processes with collaborative product lifecycle management and 3D terrestrial laser scanning tools: Reducing costs and increasing productivity* (NPS-AM-11-180; Acquisition Research Sponsored Report Series). Monterey, CA: Naval Postgraduate School.

Housel, T. J., & Bell, A. H. (2001). *Measuring and managing knowledge*. Boston, MA: McGraw–Hill/Irwin.

Housel, T., El Sawy, O., Zhong, J., & Rodgers, W. (2001, December). *Models for measuring the return on information technology: A proof of concept demonstration*. 22nd International Conference on Information Systems.

Koenig, P. C. (2009). *Real options in ship and force structure analysis: A research agenda.* Conference paper, American Society of Naval Engineers, ASNE Day (April).

Koenig, P. C., Czapiewski, P. M., & Hootman, J. C. (2008). Synthesis and analysis of future naval fleets. *Ships and Offshore Structures 3*(2), 81–89.

Komoroski, C. L. (2005). *Reducing cycle time and increasing value through the application of Knowledge Value Added methodology to the U.S. Navy shipyard planning process* (Master's thesis, Naval Postgraduate School). Monterey, CA: Naval Postgraduate School.

Komoroski, C. L., Housel, T., Hom, S., & Mun, J. (2006). *A methodology for improving the shipyard planning process: Using KVA analysis, risk simulation, and strategic real options* (NPS-AM-06-017). Monterey, CA: Naval Postgraduate School.

Laser Design. (2007). Laser scanning helps Navy save $250,000 by enabling competitive bidding [Press release].

Mun, J. (2015). *Modeling risk: Applying Monte Carlo simulation, real options analysis, forecasting, and optimization* (3rd ed.). Thomson Shore.

Mun, J. (2016). *Real options analysis: Tools and techniques* (3rd ed.). Thomson Shore.

Sturtevant, G. (2015). *Flexible ships: Affordable relevance over the ship's life cycle* [PowerPoint presentation]. Retrieved from http://www.asne-tw.org/asne/events/presentation/2015-01Sturtevant.pdf

Persistent Platforms—The DDG 51 Case

Ira Lewis—PhD, is an Associate Professor of logistics in the Graduate School of Business and Public Policy at the Naval Postgraduate School, Monterey, CA. His interests include transportation, public policy, and the international defense industry. [ialewis@nps.edu]

Abstract

In the U.S. Navy, the DDG 51 (Arleigh Burke) class of guided-missile destroyer, which first entered service in 1991, remains in production with over 70 vessels delivered. This report explores some of the key reasons for the success of this ship. The upcoming Flight III of the class, which begins procurement in fiscal year 2016, faces the challenging integration of the Air and Missile Defense Radar, which adds ballistic missile defense capability to the vessel. We conclude that the DDG 51 class features the expandability (growth margin) and open systems characteristic of a "persistent platform" that continues in production and service for a greater period of time than would have originally been contemplated.

Executive Summary

The idea behind this research originated with the Acquisition Research Program list of potential sponsored research topics for fiscal year (FY) 2015, which asked, "How do we acquire systems with performance margins and configuration flexibility to support 30 or 40 years of unknown threats?" One of the best examples of what could be called "persistent platforms" that have encountered success is the DDG 51 Arleigh Burke destroyer, a class that entered service in 1991. The DDG 51 is a multi-mission, guided-missile ship with an emphasis on air defense. The success of this class, which now includes more than 70 vessels (including the first Flight III vessel that will be procured beginning in FY2016), has been attributed to the ship's "growth margin," which is the capacity for successfully receiving new or additional equipment that is larger, heavier, or more power-intensive than in the original configuration.

The incorporation of evolving technologies into the DDG 51, culminating in a successful critical design review of Flight III's Air and Missile Defense Radar in April 2015, makes it reasonable to characterize the DDG 51 as a persistent platform. Another contributing factor is the incorporation of major systems from the DDG 1000 and LHD 6/7 vessels into Flight III, reducing acquisition risk and lowering costs. On balance, the DDG 51 represents a good example of an open system as contemplated in Department of Defense (DoD) guidance: Minimized duplication for technology development investments, and shared life-cycle costs among several major shipbuilding programs.

Critics have pointed out that Flight III lacks an Analysis of Alternatives and is proceeding on the basis of an Engineering Change Proposal rather than a new requirements and contracting process. Yet the Navy's strategy does allow for the benefits of multi-year procurement, and with production of ships split evenly between two yards using the Profit Related to Offers concept, there is stability accompanied by reasonable incentives in the industrial base—something observers of the shipbuilding industry have often expressed a desire for. With the cancellation of the CG(X) and truncation of the DDG 1000–class at three ships, the Navy will be relying on the DDG 51s for years to come to meet a wide variety of maritime requirements.

Introduction

The idea behind this paper originated with the Acquisition Research Program list of potential sponsored research topics for fiscal year (FY) 2015:

T14-031	How do we acquire systems with performance margins and configuration flexibility appropriate to support 30 or 40 years of unknown threats? What engineering standards and design protocols must be changed to accommodate non-point solution design solutions? In short how do we NOT end up with another AAAV/EFV, F-22, B-2, or other really cool, really expensive, but broadly useless system? How did we get the B-52, the LPD 4 class, and the CVN 68 class, all VERY useful and almost timeless systems that have served as first line platforms for their entire service lives and, actually beyond?

The above statement makes a legitimate point. While programs can be troubled and subject to endless delays (like the Marine Corps' Expeditionary Fighting Vehicle) or built in much smaller quantities that anticipated (e.g., F-22, B-2), these "boutique" systems can't be fairly characterized as "broadly useless." The systems portrayed as successful in the above statement share a number of qualities, including long service lives, multiple changes in mission, the ability to integrate new technologies into the prime platform, and a sufficient number of end items to generate economies of scale. The "failures," like the F-22, often procured in smaller quantities than intended, do successfully perform the relatively narrow missions they were designed for.

One of the best examples of what could be called "persistent platforms" that have encountered success is the DDG 51 Arleigh Burke guided-missile destroyer, a class that entered service in 1991. The destroyer's production run "has outlasted every other battleship, cruiser, destroyer and frigate in U.S. Navy history" (Sharp, 2009). The DDG 51 is a multi-mission destroyer with an emphasis on air defense. The success of the Arleigh Burke class, which now includes more than 70 vessels, has been attributed to the ship's "growth margin," which is the capacity for successfully receiving new or additional equipment that is larger, heavier, or more power-intensive than in the original configuration (O'Rourke, 2015a). For example, some future vessels will feature a hybrid electric drive (Scott, 2015). Additionally, modular vertical launch systems featuring the Standard Missile Three have been installed on DDG 51s (Doerry, 2012). The relatively high volume of vessels and the allocation of ships to two builders while maintaining competition (as discussed below) are also indicators of successful shipbuilding practices (Government Accountability Office [GAO], 2009).

This paper uses the "snowball" technique of data gathering. Starting with visits to the websites of DoD organizations, manufacturers, think tanks, and Congress, as well as the commercial databases available through the Naval Postgraduate School library, information is collected until redundancy begins to occur, at which point data collection ends and the analysis phase begins (Biernacki & Waldorf, 1981).

This paper begins by reviewing the background of the DDG 51 class of vessels. Second, we discuss the development of the upcoming Flight III of the vessels. The next section covers the integration challenges associated with the new Air and Missile Defense Radar aboard that series of ships. We conclude by discussing the characteristics of the DDG 51 as a persistent platform.

The Open Systems Approach

While the concept of "growth margin" is very useful for the physical capacity for renewal of a prime system, the "open systems" or "modular open systems" concepts promote an incremental or evolutionary view of the acquisition process, with each increment providing an increased level of capability. Open systems architecture enables design for affordable change, employs evolutionary acquisition and spiral development, and uses an integrated roadmap for system design and development (DoD, 2014; GAO, 2014c). Department of Defense Instruction (DoDI) 5000.02, *Operation of the Defense Acquisition System* (OUSD[AT&L], 2015a), requires the program manager to apply "open system approaches in product designs where feasible and cost effective" (enclosure 3, sec. 14).

Open system approaches form part of an overall Intellectual Property (IP) strategy according to DoDI 5000.02, Enclosure 2, Section 7b. Program needs for IP deliverables and associated license rights are considered necessary for "competitive and affordable acquisition and sustainment over the entire life cycle" (OUSD[AT&L], 2015a, enclosure 2, sec. 7b). Beyond the requirements internal to the DoD, contracts should also be structured to encourage vendors to provide open systems, and contractors must describe how they will meet modular open system requirements (Rendon, 2006; DoD, 2013, p. 5).

The following business practices are recommended by the DoD (2013) to support open systems programs:

- Seek data deliverables and rights in technical data and computer software sufficient for competition throughout the life cycle as an objective;
- Continuous competition throughout the life cycle;
- Increased capability to the warfighter on a faster development timeline;
- Reduced life cycle costs;
- Shared risks with other programs;
- Minimized duplication for technology development investments, shared life cycle costs; and
- Collaboration through peer reviews. (p. ix)

In addition to the above, an "open business model" is recommended, as such an approach "requires doing business transparently to leverage the collaborative innovation of numerous participants across the enterprise permitting shared risk, maximize asset reuse and reduce total ownership costs" (DoD, 2013, p. 139). Reuse of assets allows unique development efforts to serve several purposes, including uses that were unknown or unintended at the time of fielding of the original system. Open systems architecture is therefore a means for stimulating innovation.

A shift toward persistent platforms like the DDG 51 needs to be carried out in parallel to a move toward open systems. As stated recently in an Air Force policy document,

> To the extent that our current policies and regulations can be modified to change the paradigm from large, complex programs rife with crippling interdependencies to programs with simple, severable components, open architectures, and more distributed participation, we will enact those changes. (Department of the Air Force [DAF], 2014)

However, requirements discipline remains critical. As the GAO has found, shipbuilding should include (1) demonstrating balance among program requirements, technology demands, and cost considerations by preliminary design review, and (2) retiring technical risk and closing any remaining gaps in design requirements before a contract for detail design is awarded (GAO, 2009).

For reuse to be possible, “strategic use of data rights” is needed to moderate the monopolistic power of the vendor throughout the life cycle, which can result in “vendor lock” (DoD, 2013, pp. vii–ix, 167–176). Vendor lock gives the vendor what can become monopolistic powers, freeing the vendor to establish noncompetitive prices and become a sole source for a product or service. The DoD has “little leverage to control costs and manage performance in a vendor lock scenario” (GAO, 2014b, 2014c). As further explained by the GAO (2014c),

> The most difficult challenge is overcoming a general cultural preference within the services for acquiring proprietary systems that puts life-cycle decisions in the hands of the contractors that developed and produced those systems. Those contractors, therefore, benefit from maintaining the status quo with respect to long-term weapon system sustainment.

However, the effects of vendor lock can be mitigated through, among other practices, exploring common product lines for commonality and shared technology, rather than developing new products that correspond to individual program requirements (Wydler, 2014). Such an approach should result in modular systems and subsystems that can be openly competed.

An example of the use of an open architecture approach is the Navy’s procurement of the Ground Control Segment (GCS) for its Unmanned Aerial Vehicles (UAVs). The GCS is a common baseline for all Navy UAV programs, with users selecting mission applications from an “app store.” Within the GCS environment, all interfaces are known to all vendors, allowing a modular approach that promotes competition. A working group with representatives of more than 200 organizations defined both the interfaces and systems architecture (Lundquist, 2013).

The emphasis on reducing life-cycle costs is particularly notable. Successful, long-lived systems may have lived through the “death spiral” of decreasing quantities and increasing unit cost during their initial development and production. However, at later stages, such as following Full Operational Capability, ongoing modifications to persistent platforms may be significantly less expensive than the development of new products. The Navy’s approach to the DDG 51 class has been described as follows:

> Prior to Flight III, the program has produced three flights (I, II and IIA). Flights II and IIA included important modifications for changing mission requirements and technology updates, thus demonstrating the substantial capacity and flexibility of the base DDG 51 hull form. Flight II introduced enhanced capability in Combat Systems and Electronic Warfare. Flight IIA constituted a more significant change to the ship by incorporation of an organic dual hangar/dual helicopter aviation facility, extended transom, zonal electrical power distribution, enhanced missile capacity, and reconfigured primary radar arrays. The combined scope and means for integrating the changes for Flight III is similar to the approach used in the Flight IIA upgrade.
>
> The previous ship system changes were successfully executed by ECPs [Engineering Change Proposals] introduced via the existing systems engineering processes on both Flight II and IIA in support of the ongoing construction program. This methodology takes advantage of Navy and prime contractor experience with the proven processes while offering effective and

> efficient introduction of the desired configuration changes. It also provides the more affordable and effective approach toward producing this enhanced ship capability in lieu of starting a new ship design to incorporate the same capabilities into a new production line for ship construction. (ASN[RDA], 2015)

Another relevant example is the development of the F/A-18 E/F Super Hornet into the EA-18G Growler. It would have cost far more to develop a new aircraft solely to meet the electronic attack mission. It is interesting that the Royal Australian Air Force has ordered that 12 of its acquisition of 24 Super Hornets be wired for potential conversion to Growlers, should the need arise (Fulghum, 2011). The success of the Super Hornet has been aptly described as follows:

> Although the E/F version continues the basic name and design concept of the original F/A-18, it was significantly redesigned. While originally maintaining 90% avionics and software commonality with the F/A-18C/D model, its airframe is 25% larger, and features radar, avionics, and weapons upgrades and more powerful engines. Its weapons and fuel stores capacity have been significantly increased, and it can be utilized as an aerial refueling tanker. The newer models also provide frontal stealth qualities. The enhanced capabilities of the F/A-18E/F are a possible explanation for the Navy's decision not to seek to develop a direct replacement for the F-14 Tomcat. The multiple mission suites of the Hornet and Super Hornet may have allowed the retirement of a sizeable number of specialized Navy aircraft which had been fulfilling its combat aircraft roles with an associated reduction in logistics complexity. (Franck, Lewis, & Udis, 2011)

A final note on the F/A-18 E/F is from the perspective of industrial policy. Only Lockheed Martin and Boeing remain as U.S. manufacturers of fighter aircraft. If production of the Super Hornet ends without a replacement aircraft, then Lockheed Martin will have a monopoly. There are some practical reasons to keep the Super Hornet in production, such as the fact that it can act as a tactical refueling aircraft for other F/A-18s, a capability the F-35 does not have. The Navy has a replacement program for the Super Hornets known as the F/A-XX, but that acquisition is far off. The future of Boeing's vast St. Louis plant, which currently produces the F-15 in small quantities for export customers as well as the F/A-18 E/F, depends in great part on whether the DoD wishes there to be one or two providers of combat aircraft. As we discuss shortly, industrial base considerations play a more pronounced role in shipbuilding policy than is the case for aircraft manufacturing.

Our brief discussion of a number of weapon systems has emphasized the benefits of continuing production of existing platforms that are flexible enough to meet the challenges of changing requirements and evolving technologies. We now turn to a review of the features of the DDG 51 vessels.

DDG 51 Class: Background

The DDG 51 (Arleigh Burke) class of multi-mission destroyer is unique because of the size of the class, with 72 ships having been procured from FY1985 through FY2015. The DDG 51 class shares the Aegis defense system with the Ticonderoga (CG 47) class of cruiser, and both types of vessels are often referred to generically as "Aegis ships." The original main role of the Arleigh Burke class was air defense and mid-ocean operations, but the destroyers are now being outfitted for ballistic missile defense operations (McCullough, 2013; O'Rourke, 2015c).

The first 28 DDG 51s (DDG 51–DDG 78) are referred to as Flights I/II. DDG 79–123 are part of Flight IIA, which notably includes a hangar for two embarked Light Airborne Multi-Purpose System MK-III SH-60R helicopters omitted in Flight I/II ships (O'Rourke, 2015a; Program Executive Officer Ships, 2015). Current plans are to build at least 22 Flight III vessels (Director, Operational Test and Evaluation, 2015). Additionally, in-service vessels of the DDG 51 class are undergoing a DDG Modernization Program that will provide mid-life upgrades (Program Executive Officer Ships, 2015).

No DDG 51s were procured during FY2006–FY2009 (O'Rourke, 2015c). Credit for the restart of production of the Arleigh Burke class in 2011 must be attributed in part to the significant development problems encountered with the Zumwalt-class (DDG 1000) destroyer. Only three of these vessels are under construction, with further production of DDG 51s in effect replacing the troubled Zumwalts. The DDG 1000s were designed with a number of revolutionary features, including a stealthy tumbledown hull and electric-drive propulsion. Of the 10 transformational technologies incorporated in the DDG 1000, four were found to be immature at Milestone B. The decision to terminate the DDG 1000 program and restart the DDG 51 has been described as follows:

> The FY 2011 budget decision to truncate the DDG-1000 program at three ships and restart DDG-51 production was largely due to a change in the perceived threat and mission priority by Navy senior leadership. Priority was placed on ballistic missile defense rather than the original DDG-1000 precision and volume fire support mission.
>
> The radar hull study recommended the DDG-51 hull form with a new advanced missile defense radar (AMDR) as more effective in the ballistic missile defense mission than DDG-1000. DDG-51 with AMDR was also assessed to have less cost risk. (Blickstein et al., 2011)

The Navy has also cancelled its CG(X) cruiser development program, with the DDG 51 intended as a bridge in capability, particularly in air and ballistic missile defense, to the eventual acquisition of new cruisers (Bliss, 2010; GAO, 2012; Hagerty, Stevens, & Wolfe, 2008). It could be suggested that the difficulties encountered by the DDG 1000 and CG(X) have proven to be advantageous for continued production of the DDG 51.

Traditional Navy policy has been to keep surface combatants such as destroyers and cruisers in service for 25 years. However, current projections show that the Navy intends to maintain the DDG 51 Flights IIA and III in service for 40 years. The Congressional Budget Office (CBO) viewed this plan with skepticism, pointing out that it might not be cost-effective or technically feasible to maintain and refit these complex vessels for such a long period of time (*Analysis of the Navy's*, 2013; CBO, 2013).

In a unique arrangement, construction of the class is split between General Dynamics' Bath Iron Works (Bath, ME) and Huntington Ingalls Industries' Ingalls Shipbuilding (Pascagoula, MS). The Navy allocates hulls equally to both yards, and the yard that submits a bid lower than what had been assigned by the Navy receives a higher profit margin. This approach, termed Profit Related to Offers bidding (a variant of Fixed Price Incentive [Firm Target]), has been described as competition for profit rather than for quantity and is considered a successful means of dealing with the challenge of maintaining competition in the face of small procurement quantities (*Case Studies*, 2014; GAO, 2012; Kendall, 2015; OUSD[AT&L], 2014a), and has won a David Packard Excellence in Acquisition Award (Freedberg, 2012). As emphasized by the GAO (2009),

> Moving to fixed-price contracting is an important element in changing the paradigm for shipbuilding programs—fixed-price contracting can only be used if risk is appropriately retired by the time a contract for construction is agreed on and a clear understanding of the effort needed to deliver the ship exists.

Further stability is offered by the DoD's funding strategy, which proposes buying two DDG 51s annually for FY2015 through FY2019 (Pentagon Budget, 2014). This approach should assist both builders in recruitment and retention of skilled workers (Arena et al., 2006, pp. 61–62; GAO, 2009). It has been pointed out that industrial base considerations play an explicit role when developing a ship program's acquisition strategy (Drezner et al., 2011, p. 39).

Flight III Development

The development of the most recent version of the ship, known as Flight III, has been a challenge for the Navy, yet confirms the presence of the "growth margin" discussed previously. The Navy plans to begin procurement of 33 Flight III ships in FY2016, with the first vessel, DDG 124, expected to achieve initial operating capability in 2023 (Scott, 2015). Similarly, it has been observed that modifying the DDG-51

> over time has used up some of the design's growth margin. The Flight III DDG-51 would in some respects have less of a growth margin than what the Navy would aim to include in a new destroyer design of about the same size. (O'Rourke, 2014)

The current development challenges have been described as follows:

> DDG 51 Flight III ships are expected to feature new electric plants, new air-conditioning plants, and the AMDR. According to the Navy, the new electric plants are based on a design used on DDG 1000 and modification will be required for integration with DDG 51. The DDG 1000 electrical system has faced delays in completing testing. Detail design work for Flight III will begin at the end of fiscal year 2014, according to the Navy. Adding AMDR to DDG 51 will result in a significant redesign of the ship and the Navy expects that Flight III will result in changes to more than 25 percent of Flight IIA drawings, although the Navy believes many of these will be minor alterations. The Navy will need AMDR's design assumptions, such as its size, shape, weight, and power and cooling requirements in order to accurately redesign the ship. However, the Navy only recently awarded a contract for AMDR system development and the AMDR program is at least 6 months behind schedule. Based on its current schedule, the Navy plans to begin detail design work for Flight III at the end of fiscal year 2014—before AMDR has demonstrated full maturity-adding risk and uncertainty to the DDG 51 program. (GAO, 2014a).

It should be noted, however, that the AMDR successfully completed its Critical Design Review in April 2015, including "a thorough review of all design information to ensure the system will meet required specifications within cost and schedule constraints" (Department of the Navy [DoN], 2015) Among the benefits of the AMDR, other than its contribution to missile defense, will be a significant reduction in maintenance, allowing crews to spend more time actually operating the radar (Taylor, 2014). The AMDR is the largest new system on the ship, which otherwise uses existing systems from Flight IIA and some adopted in other classes of vessel:

> According to Capt [Mark] Vandroff [DDG 51 Shipbuilding Program Manager], development risk has been kept in check by the widespread re-use of existing

> machinery and systems. “When I needed electrical power, my friends in PMS 320 [NAVSEA’s Electric Ship Office], they said ‘Yes, we’ve got a generator for you. We’ll get you a power conversion module. They’re already out there, you don’t have to invent something new.’”
>
> “That reduces our R&D [research and development] cost. You take a look at my budget, you will see that we are really tightening our purse strings for so much procurement. That’s always been our philosophy on Flight III. We’ve got one new thing we are inventing [AMDR]; everything else is what’s out there.”
>
> “So I take the DDG 1000 generators, I take the LHD 6/7 electrical distribution system, I marry them up and I give the ship more power. I need to convert that power to 1000 V DC, so I take a power conversion module now under contract with DRS. We’re going with something we understand.” (Scott, 2015)

As just stated, Flight III will use the generator (and power conversion module) from the DDG 1000, and the 4160 VAC Electric Plant from the LHD 6/7 class (ASN [RDA], 2015). Sharing the risk of systems from other programs and reusing existing systems were described as being among the characteristics of an open system approach. The reuse of systems from other classes of vessels is also an effective means of dealing with the “vendor lock” and IP challenges discussed previously (Wydler, 2014). In constant FY2013/2014 dollars, the cost of each Flight III vessel has been estimated by the DDG 51 program manager at $1.7 billion for a total fleet of 22 destroyers with two ships procured annually. In comparison, Flight IIA vessels cost an average of $1.5 billion each (Jean, 2014).

The potential lack of sufficient growth margin in the DDG 51 Flight III, combined with the small number of DDG 1000s to be built, may raise questions as to the long-term capabilities of the combatant fleet. For example, electric drive, a feature of the DDG 1000s but absent from the DDG 51 class, is intended in part to provide sufficient power for future weapons such as lasers and the AMDR. One suggestion has been to lengthen the hull of the ships to permit future insertion of technologies such as electric drive and more advanced versions of the AMDR (DoN, 2014; GAO, 2012; Fabey, 2012; O’Rourke, 2015a). It should be noted that the DDG 51 does not represent a full replacement of the projected capabilities of the three-ship DDG 1000 class and the cancelled CG(X):

> The Navy’s pre-2008 plan to procure DDG-1000 destroyers and then CG(X) cruisers based on the DDG-1000 hull design represented the Navy’s roadmap at the time for restoring growth margins, and for introducing into the cruiser-destroyer force significant numbers of ships with integrated electric drive systems and technologies for substantially reducing ship crew sizes. The ending of the DDG-1000 and CG(X) programs in favor of continued procurement of DDG-51s leaves the Navy without an announced roadmap to do these things, because the Flight III DDG-51 will not feature a fully restored growth margin, will not be equipped with an integrated electric drive system or other technologies that could provide ample electrical power for supporting future electrically powered weapons, and will not incorporate features for substantially reducing ship crew size or for otherwise reducing ship O&S [operations & support] costs substantially below that of Flight IIA DDG-51s. (O’Rourke, 2015a)

However, the Navy still plans to include electric drive in at least some of the Flight III ships, although space constraints are currently a concern. There are also plans to retrofit 36

Flight IIA destroyers with hybrid electric drive (Mazumdar, 2012), with a trial on one vessel described as follows:

> The fuel savings offered by a hybrid electric drive have prompted consideration of back-fitting a CODLOG [combined diesel electric or gas] option to the USN's most numerous class, the Arleigh Burke destroyers. Changing the propulsion configuration of a warship post-build potentially raises a large number of engineering challenges which, if costly, would negate the savings to be derived from reduced fuel usage. A pilot installation of a 1.9 MW electric motor attached to an existing lay shaft of the DDG 51 gearbox was trialled in USS *Truxtun* in 2012 [Flight IIA], and is scheduled for a much wider roll-out from 2017. Electrical power is drawn from the main electrical distribution system, which in the DDG 51 is provided by up to three 2.5 MW gas turbine-powered generators. The configuration only provides electric drive to the port shaft, a compromise driven by cost versus benefit. (Scott, 2015)

The GAO (2012) has also pointed out that adding a hybrid electric drive to the ship in Flight III "would require additional design changes to accommodate the new motors and supporting equipment." Electric drive will be particularly important should the Navy proceed with plans to install laser weapons on the Flight III vessels (O'Rourke, 2015b).

A GAO review of plans for the DDG 51 noted that the Navy's Radar/Hull Study of 2009, which recommended the integration of the AMDR onboard Flight III vessels, "does not provide an adequate evaluation of combat system and ship characteristics, and does not include key elements that are expected in an AOA [analysis of alternatives] that would help support a sound, long term acquisition program decision" (GAO, 2012). These elements include computer processing ability, cyber warfare capability, reliability, information assurance capability, usability, proprietary versus open architecture combat systems, and scalability (GAO, 2012) .

A related concern is that the Navy is considering a scaled-down (12-foot) version of the AMDR for Flight III; however, this smaller radar may not be able to meet the Navy's Integrated Air and Missile Defense (IAMD) requirements. This challenge brings into question the entire strategy of using the DDG 51 as a means to bring ballistic missile defense to Aegis vessels (Fabey, 2011). The Navy is, however, considering an as-yet-unproven concept of networking IAMD capabilities, which may reduce the requirements for the AMDR carried onboard each vessel. The dilemma was summarized as follows:

> Analysts and contractors say it's starting to appear that the Flight IIIs will be heavily modified to accommodate the AMDR. But the Navy's top shipbuilder executive warns against that course. "Sometimes we get caught up in the glamour of the high technology," [Huntington Ingalls Industries CEO Mike] Petters says. "The radars get bounced around. They get changed. Their missions get changed. … The challenge is if you let the radars drive the ships, you might not get any ships built." (Fabey, 2011)

It should be noted, as mentioned in the next section, that the AMDR will now be 14 feet in width, so concerns about a scaled-down radar were premature (ASN[RDA], 2015).

Integration Challenges

The ship is already packed with equipment, with the DDG 51 being described as the "densest surface combatant class" (GAO, 2012). In a moment of pique, the CBO referred to a potential DDG 51 Flight IV as the DDG(X), stating that the CBO "considers it unlikely that the Navy would or could use the DDG-51 design for the next-generation destroyer" (Analysis of Navy's Plan, 2013). However, the point is now moot, as Flight IV was cancelled in 2014 in order to fund replacement of the Ohio-class ballistic missile submarine (LaGrone, 2014). One observer described the Flight III challenge as follows:

> Once this ship is built—mostly by performing superhuman feats of engineering on an already crowded ship—there will be a limited margin (power, weight, cooling, displacement) for future upgrade. What the Navy buys is essentially what it will have for the life of the ship, with modest upgrades available through computer program refreshes and the like, but not the kind of upgrades available (theoretically) in a more modular design in which power, cooling, radars, and weapons, etc. are treated as commodities that can be swapped out when increased capability (payloads) is available. (McGrath, 2013)

In a similar vein, the GAO has assessed that the Navy's acquisition approach to Flight III is not commensurate with risks, both financial and technical. For example, multi-year procurement is being used despite the high level of uncertainty associated with a challenging ship redesign:

> Further, technical studies about Flight III and the equipment it will carry are still underway, and key decisions about the ship have not yet been made. DDG 123 [the last Flight IIA ship] is not due to start construction until fiscal year 2016. If the Navy proceeds with this plan it would ultimately be awarding a multiyear contract including this ship next fiscal year, even though design work has not yet started and without sufficient knowledge about cost or any construction history on which to base its costs, while waiting until this work is done could result in a more realistic understanding of costs. Our prior work has shown that construction of lead ships is challenging, the risk of cost growth is high, and having sufficient construction knowledge is important before awarding shipbuilding contracts. (GAO, 2012)

Despite the Navy's optimism, there is reason for some concern regarding acquisition plans for up to 43 DDG 51 vessels. Flights I through IIA were highly successful during both acquisition and operation. However, attempts to retrofit the vessel with the AMDR, increased power and cooling requirements, and a high density of equipment are driving a significant increase in the risk of this "proven" Aegis destroyer. The AMDR requires five times more power and 10 times more cooling capacity than the SPY-1D(V) radar being replaced (LaGrone, 2012). A recent discussion, however, showed that the Navy is aware of the risk and is managing accordingly:

> Three mantras characterise the Flight III design and development effort: minimum change, minimum risk, and maximum re-use.
>
> "We know what we need to do," said Capt [Mark] Vandroff [DDG 51 Shipbuilding Program Manager]. "We need to get a SPY +15 dB radar onto a DDG 51 hull and deliver it to the fleet. Every other one of those requirements after that, some of them might be nice but we might say, 'I'm not taking that stuff today. I'm going to do something else. I'm going to keep the risk low.' So the only technology that's getting on [Flight III] is something that's already ripe and ready right now." (Scott, 2015)

With the DDG 1000 class reduced to a boutique fleet of three vessels and the cancellation of the CG(X) program, the burden of IAMD rests squarely on the DDG 51. A key finding of the Radar/Hull study was that the DDG 51 design represented better value for money than the DDG 1000 for delivering IAMD capability (Scott, 2015). The GAO (2012) has recommended a “robust operational test program” for Flight III, despite Navy reticence. While there remains no formal AOA for Flight III, the DoD position is that documents such as the Radar/Hull study, when assembled, constitute an equivalent body of knowledge (GAO, 2012). The GAO has commented on the maturity and testing of the AMDR as follows:

> All four of AMDR’s critical technologies—digital-beam-forming; transmit-receive modules; software; and digital receivers/exciters—are approaching full maturity, and program officials state that AMDR is on pace to meet DDG 51 Flight III’s schedule requirements. In 2015, the contractor is expected to complete an engineering development model consisting of a single full-sized 14 foot radar array—as opposed to the final four array configuration planned for installation on DDG 51 Flight III—and begin testing in the contractor’s indoor facilities. Following the critical design review, scheduled for April 2015, the program plans to install the array in the Navy’s land-based radar test facility in Hawaii for further testing in a more representative environment. However, the Navy has no plans to test AMDR in a realistic (at-sea) environment prior to installation on the lead DDG 51 Flight III ship. Though the Navy is taking some risk reduction measures, there are only 15 months planned to install and test the AMDR prototype prior to making a production decision. Delays may cause compounding effects on testing of upgrades to the Aegis combat system since the Navy plans to use the AMDR engineering development model in combat system integration and testing. (GAO, 2015)

Furthermore, concerns have been raised about the lack of a self-defense test ship in Flight III’s operational test and evaluation program, as well as the absence of the at-sea testing plan with the AMDR, the testing plan having been specifically mandated by the Deputy Secretary of Defense in a March 6, 2014, memorandum (Director, Operational Test and Evaluation, 2015). The requirement for an at-sea testing plan is echoed by the GAO’s (2009) recommendations on shipbuilding practice, which state that “critical technologies be developed into representative prototypes and successfully demonstrated in a relevant environment.”

In a similar vein, a Congressional Research Service report expressed skepticism that the Flight III design potentially precludes fitting the vessel with a high-power laser due to lack of sufficient electrical power or cooling capacity. Lasers were viewed as being essential to the Navy’s ability to counter anti-ship cruise missiles or ballistic missiles in the future (O’Rourke, 2015a).

Additionally, the GAO has expressed concern that the Navy plans to use ECPs to the existing Flight IIA multiyear procurement contracts, rather than establish new contracts, to construct the first three Flight III ships (GAO, 2015). Focusing on the AMDR as the major new system aboard Flight III, Congress mandated that the Navy submit a report on the AMDR ECP before going ahead with the work. That report explained the Navy’s strategy for Flight III development:

> ECP development is a fundamental systems engineering approach; an approach currently implemented in the DDG 51 program that has been continuously updated and improved since the program’s inception in the early 1980s and has resulted in the successful delivery of 62 DDG 51 Class destroyers. The last three ships of the FY13-17 MYP [Multi-Year Procurement] are designated as Flight III beginning with one of the FY16

ships. The Flight III is a modified repeat of the existing baseline and will be centered on the addition of an IAMD capability in the form of the AMDR-S [S-band radar], associated enhanced combat systems elements and requisite supporting HM&E [Hull, Mechanical, and Electrical] changes. These changes will be incorporated via discrete ECPs with the same proven processes and rigor that produced successful Flight II and IIA upgrades to the class. (ASN [RDA], 2015)

The Navy's approach to Flight III balances the incremental need for ballistic missile defense (which led to the AMDR) with the demanding systems integration challenges of a complex vessel. The overall approach is somewhat conservative in the sense of limiting changes to those required by the AMDR itself or the supporting systems. In doing so, the goal is to continue the incremental approach that has characterized the successful development of the DDG 51 class.

Conclusion: A Persistent Platform?

With more than 70 vessels delivered or on order, the large size of the DDG 51 class of guided-missile destroyer leads to the reasonable finding that this ship has been successful. The technology insertion and myriad of changes made to the ship over its four flights (I, II, IIA, and III), while retaining the same hull and essentially the same turbine, are further confirmation of the endurance of the Arleigh Burke's fundamental design. Had funding not been cut to fund the Ohio-class submarine replacement, there might have been a Flight IV as well.

The successful incorporation of evolving technologies into the class, culminating in a successful critical design review of Flight III's AMDR in April 2015, makes it reasonable to characterize the DDG 51 as a persistent platform. Another supporting factor is the incorporation of major systems from the DDG 1000 and LHD 6/7 vessels into Flight III, reducing acquisition risk and lowering costs. The GAO's harsh January 2012 criticism of the planned insertion of the AMDR was perhaps justified; the watchdog agency's concerns were probably the key factors behind Congress mandating a report on the AMDR ECP, which was released in February 2015. On balance, the DDG 51 does represent a good example of an open system as contemplated in DoD guidance: minimized duplication for technology development investments, as well as shared life-cycle costs among several major shipbuilding programs.

Critics have pointed out that Flight III lacks an AOA and is proceeding on the basis of an ECP rather than a new requirements and contracting process. Yet the Navy's strategy does allow for the benefits of multi-year procurement, and with production of ships split evenly between two yards using the Profit Related to Offers concept, there is stability accompanied by reasonable incentives in the industrial base—something observers of the shipbuilding industry have often expressed a desire for.

The DDG 51 has become a persistent platform because it is adaptable and was built with sufficient growth margin to accommodate evolving operational needs such as ballistic missile defense. With the cancellation of the CG(X) and truncation of the DDG 1000 class at three ships, the Navy will be relying on the DDG 51s for years to come to meet a wide variety of maritime requirements.

References

An analysis of the Navy's fiscal year 2014 shipbuilding plan: Hearing before the Subcommittee on Seapower and Projection Forces, Committee on Armed Services, *House of Representatives,* 113th Cong. (October 23, 2013) (testimony of E. J. Labs). Retrieved from http://www.cbo.gov

Arena, M. V., Blickstein, I., Younossi, O., & Grammich, C. A. (2006). *Why has the cost of Navy ships risen? A macroscopic examination of the trends in U.S. Naval ship costs over the past several decades* (MG-484). Santa Monica, CA: RAND. Retrieved from http://www.rand.org

Arleigh Burke (Flight IIA) class. (2015, April 2). *Jane's Fighting Ships*. Retrieved from http://janes.ihs.com

Assistant Secretary of the Navy (Research, Development, and Acquisition; ASN [RDA]). (2015, February). *Report to Congress: DDG 51 Flight III ships air and missile defense radar engineering change proposal*. Retrieved from http://insidedefense.com

Biernacki, P., & Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological Methods & Research, 10*(2), 141–163.

Blickstein, I., Boito, M., Drezner, J. A., Dryden, J., Horn, K., Kallimani, J. G., … Wong, C. (2011). *Root cause analysis of Nunn-McCurdy breaches (Vol. 1)* (MG-1171/1-OSD). Santa Monica, CA: RAND. Retrieved from http://www.rand.org

Bliss, G. R. (2010, May 21). *PARCA root cause analysis for the DDG-1000 program* [Memorandum for Under Secretary of Defense for Acquisition, Technology and Logistics from Director, Performance Assessments and Root Cause Analyses].

Case studies in DoD acquisition: Finding what works: Statement before the House Armed Services Committee, House of Representatives, 113th Cong. (June 24, 2014) (testimony of Ronald O'Rourke). Retrieved from http://armedservices.house.gov

Congressional Budget Office. (2013, October). *An analysis of the Navy's fiscal year 2014 shipbuilding plan*. Retrieved from http://www.cbo.gov

Department of the Air Force (DAF). (2014, July). *America's Air Force: A call to the future*. Retrieved from http://www.af.mil

DoD. (2013, June). *DoD open systems architecture contract guidebook for program managers* (Ver. 1.1). Retrieved from http://www.acq.osd.mil

DoD. (2014, June 23). *Office of the Deputy Assistant Secretary of Defense (Systems Engineering): Initiatives*. Retrieved from http://www.acq.osd.mil

Department of the Navy (DoN). (2014, June*). Report to Congress on the annual long-range plan for construction of Naval vessels for FY2015*. Washington, DC: Deputy Chief of Naval Operations (Integration of Capability and Resources [N8]).

Department of the Navy (DoN). (2015, May 15). *Navy's Air and Missile Defense Radar (AMDR) completes system critical design review* (U.S. Navy Release NNS150515-20). Retrieved from http://www.navy.mil/submit/display.asp?story_id=87139

Director, Operational Test and Evaluation. (2015, January). *FY 2014 annual report*. Retrieved from http://www.dote.osd.mil

Doerry, N. H. (2012). Institutionalizing modular adaptable ship technologies. *Society of Naval Architects and Marine Engineers Annual Meeting and Expo 2012* (Paper No. SNAME-047-2012). Retrieved from http://www.sname.org

Drezner, J. A., Arena, M. V., McKernan, M., Murphy, M., & Riposo, J. (2011). Are ships different? *Policies and procedures for the acquisition of ship programs* (MG-991-OSD/NAVY). Santa Monica, CA: RAND. Retrieved from http://www.rand.org

Fabey, M. (2011, July). Flight of fancy. *Defense Technology International, 5*(7), 19. Retrieved from EBSCOHost.

Fabey, M. (2012, April). Flight fight. *Defense Technology International, 6*(4), 36. Retrieved from EBSCOHost.

Franck, R. E., Lewis, I., & Udis, B. (2011, April 30). International perspectives of the impact of export control and technology transfer regimes: The F/A-18 case study. In *Proceedings of the Eighth Annual Acquisition Research Symposium* (Vol. 1). Monterey, CA: Naval Postgraduate School. Retrieved from http://www.acquisitionresearch.net

Freedberg, S. J. (2012, November 12). Can Navy afford next-gen DDG-51 destroyer, Packard Award or not? *Breaking Defense*. Retrieved from http://breakingdefense.com

Fulghum, D. A. (2011, February 21). Aardvark to Super Hornet. *Aviation Week & Space Technology, 173*(7), 60.

GAO. (2009, May). *Best practices: High levels of knowledge at key points differentiate commercial shipbuilding from Navy shipbuilding* (GAO-09-322). Retrieved from http://www.gao.gov

GAO. (2012, January). *Arleigh Burke destroyers: Additional analysis and oversight required to support the Navy's future surface combatant plans* (GAO-12-113). Retrieved from http://www.gao.gov

GAO. (2014a, March). *Defense acquisitions: Assessments of selected weapon programs* (GAO-14-340SP). Retrieved from http://www.gao.gov

GAO. (2014b, May). *Defense contracting: Early attention in the acquisition process needed to enhance competition* (GAO-14-395). Retrieved from http://www.gao.gov

GAO. (2014c, June 26). *Defense acquisitions: Review of private industry and Department of Defense open systems experiences* (GAO-14-617R). Retrieved from http://www.gao.gov

GAO. (2015, March). *Defense acquisitions: Assessments of selected weapon programs* (GAO-15-342SP). Retrieved from http://www.gao.gov

Hagerty, J. C., Stevens, P. T., & Wolfe, B. D. (2008, December). *DDG 1000 vs. DDG 51: An analysis of U.S. Navy destroyer procurement* (MBA Professional Report). Monterey, CA: Naval Postgraduate School. Retrieved from http://calhoun.nps.edu

Jean, G. (2014, April 8). Navy League 2014: USN Arleigh Burke Flight IIs to feature wider stern. *Jane's Navy International*. Retrieved from http://janes.ihs.com

Kendall, F. (2015, May–June). Tying profit to performance: A valuable tool, but use with good judgment. *Defense AT&L*. Retrieved from http://www.dau.mil/pubscats/Pages/DefenseAtl.aspx

LaGrone, S. (2012, February 3). Analysis: Arleigh Burke update costs and challenges mount up. *Jane's Defence Weekly*. Retrieved from http://janes.ihs.com

LaGrone, S. (2014, July 15). Navy cancelled new destroyer flight due to Ohio replacement submarine costs. *USNI News*. Retrieved from http://www.usni.org

Lundquist, E. (2013, November). Open business model. *Seapower, 56*(11), 44–47.

Mazumdar, M. (2012, February 6). Arleigh-Burke gets HED start in fuel efficiency race. *Jane's Navy International*. Retrieved from http://janes.ihs.com

McCullough, B. (2013, January). The right destroyer at the right time. *United States Naval Institute Proceedings, 139*(1). Retrieved from ProQuest.

McGrath, B. (2013, March 26). Opinion: Navy should avoid a Flight III Arleigh Burke. *USNI News*. Retrieved from http://www.usni.org

Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics (OUSD[AT&L]). (2014a, September 19). *Better Buying Power 3.0* [White paper]. Retrieved from http://www.acq.osd.mil

Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics (OUSD [AT&L]). (2014b, December). *Selected acquisition report: DDG 51 Arleigh Burke class guided missile destroyer*. Washington, DC: Author.

Office of the Under Secretary of Defense for Acquisition, Technology, & Logistics (OUSD[AT&L]). (2015a, January 7). *Operation of the defense acquisition system* (DoD Instruction 5000.02). Washington, DC: Author.

Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics (OUSD[AT&L]). (2015b, April 9). *Better Buying Power 3.0 implementing guidance*. Retrieved from http://www.acq.osd.mil

O'Rourke, R. (2014, June 4). *Navy DDG-51 and DDG-1000 destroyer programs: Background and issues for Congress* (RL32109). Washington, DC: Congressional Research Service. Retrieved from http://www.fas.org

O'Rourke, R. (2015a, March 24). *Navy DDG-51 and DDG-1000 destroyer programs: Background and issues for Congress* (RL32109). Washington, DC: Congressional Research Service. Retrieved from http://www.fas.org

O'Rourke, R. (2015b, June 1). *Navy shipboard lasers for surface, air and missile defense: Background and issues for Congress* (R41526). Washington, DC: Congressional Research Service. Retrieved from http://www.fas.org

O'Rourke, R. (2015c, June 9). *Navy Aegis Ballistic Missile Defense (BMD) program: Background and issues for Congress* (RL33745). Washington, DC: Congressional Research Service. Retrieved from http://www.fas.org

Pentagon budget 2015: US Navy prioritises shipbuilding at expense of aviation procurement. (2014, March 5). *Jane's Defence Weekly*. Retrieved from http://janes.ihs.com

Program Executive Officer Ships. (2015, June 10). *DDG 51 program summary*. Retrieved from http://www.navsea.navy.mil/teamships/PEOS_DDG51/default.aspx

Rendon, R. (2006, April 30). Using a modular open systems approach in defense acquisitions: Implications for the contracting process. In *Proceedings of the Third Annual Acquisition Research Symposium* (NPS-CM-06-069). Monterey, CA: Naval Postgraduate School. Retrieved from http://www.acquisitionresearch.net

Scott, R. (2015, May 20). Beneath the skin: U.S. Navy DDG 51 Flight III guided missile destroyer. *Jane's Defence Weekly, 52*(27). Retrieved from ProQuest.

Sharp, D. (2009, December 31). After 2-plus decades, Navy destroyer breaks record. *San Diego Union-Tribune.* Retrieved from http://www.utsandiego.com

Shipbuilding Plan: Congressional Budget Office testimony before the Subcommittee on Seapower and Projection Forces, Committee on Armed Services, U.S. House of Representatives. Retrieved from http://www.cbo.gov

Taylor, D. P. (2014, November). A more operational DDG. *Seapower*, *57*(9), 26–27.

Wydler, V. L. (2014). Gaining leverage over vendor lock to improve acquisition performance and cost efficiencies. In P*roceedings of the 11th Annual Acquisition Research Symposium* (Vol. 1; pp. 132–147). Retrieved from http://www.acquisitionresearch.net

Applying Principles of Set-Based Design to Improve Ship Acquisition

Eric Rebentisch—is a Research Associate at MIT's Sociotechnical Systems Research Center. He leads the Center's Consortium for Engineering Program Excellence and previously led the Lean Advancement Initiative's Enterprise Product Development research. Dr. Rebentisch's portfolio of research projects include studies of the integration of program management and systems engineering and performance benchmarking of the U.S. shipbuilding industry. [erebenti@mit.edu]

CDR John Genta, USN—graduated in June 2016 from the Massachusetts Institute of Technology with master's degrees in Naval Engineering and System Design and Management. He is assigned as an engineering duty officer at the Portsmouth Naval Shipyard. [john.a.genta@navy.mil]

Abstract

Set-based design (SBD) is a relatively new complex product development method. Its use has been well-researched in the automotive industry and to a lesser extent in other industries, and although it requires an upfront investment in resources, it has been shown to reduce design cycle-time, later stage re-work, and total ownership cost, and to improve design knowledge capture. Since 2005, the U.S. Navy has self-identified ship design as a process improvement priority and embarked in design tool and policy changes which resulted in the "Two Pass/Six Gate" process in 2008. Subsequent U.S. Navy ship design and acquisition actions have presented an opportunity to research and analyze the amenability of SBD, and its proposed benefits, with the U.S. Navy's Two Pass/Six Gate process to realize the efficiencies sought by acquisition executives. This study explored the application and benefits of using set-based design in acquisition programs. It identified specific changes to the existing Two Pass/Six Gate process in order to enable more widespread use of set-based design to improve the outcomes of complex acquisition programs.

Introduction

The 2005 National Shipbuilding Research Program (NSRP) Strategic Investment Plan (SIP) stated that ship design was the number one factor contributing to increased ship construction costs, and in 2007, the Commander of Naval Sea Systems Command (NAVSEA) was quoted as saying the U.S. Navy (USN) needs to re-establish its roots in terms of disciplined ship design (Keane, Firemann, et al., 2009; Sullivan, 2008). Since 2005, the USN has identified ship design processes and tools as the main problems leading to unaffordable ships.

The USN has explored using a product design approach known as Set-Based Design (SBD) or Set-Based Concurrent Engineering (SBCE). SBD, as a philosophy, has been utilized by Toyota Motor Corporation (TMC) to achieve acclaimed automobile manufacturing dominance. Ultimately, it aided TMC in producing better cars faster than its competitors. SBD has been shown to reduce product development cycle-time and has been touted as a contributing reason for TMC's dominance in the late 20th century (Ward et al., 1995). Further research of SBD use in other manufacturing industries has shown products designed via SBD result in reduced production cost (Raudberget, 2010). Producing better ships faster and cheaper is a process the USN desires to emulate by using SBD. Concurrent with sampling SBD, the USN has produced a suite of design tools to align with this new method of ship product development (Kassel, Cooper, & Mackenna, 2010).

This paper sets out to explore how this new approach to design and associated processes and tools might be utilized inside the SECNAV 5000 acquisition instructions and within the confines of the DoD/JCIDS/PPBE socio-technological system to realize efficiency gains in USN ship design and acquisition.

Different Approaches to Ship Product Development

SBD is contrasted with traditional or point-based design in a number of ways that are described in the following sections.

Point-Based Design (PBD)

The typical approach to design begins by defining a problem and then generating many alternative solutions (Chapman, Bahill, & Wymore, 1992). After some preliminary analysis, engineers select the alternative that appears to be the best, and then analyze, evaluate, and modify it until a satisfactory solution emerges (Ward et al., 1995). If all the initial alternative solutions could be graphed, engineers would know exactly which "point" in the design space they are analyzing, evaluating, and modifying. This approach of selecting a specific point in the design space and optimizing it is referred to as "point-based" design. However, with a point-based design, often as the fidelity of the analyses increases, design flaws begin to surface that require quick solutions to bring the design back into the feasible solution space. Often the design cannot be altered enough to achieve a feasible solution, at which point a new design alternative is chosen to re-start the design. The primary attribute of this approach is that a single solution is synthesized first, then analyzed and changed accordingly (Liker et al., 1996). A PBD process can be summarized by the following five steps (Bernstein, 1998; Liker et al., 1996):

1. Research the problem. During this step, designers inquire with the customer to clearly set problem requirements.
2. Once the requirements are known, engineers and designers use experience to quickly determine a large variety of potential solutions.
3. Engineers then perform preliminary analysis on all alternatives to determine a single, feasible, most opportunistic solution for further analysis.
4. The chosen concept is then analyzed and modified in detail to achieve all product requirements established in Step 1.
5. If the detailed design cannot be modified to meet all requirements, the process starts over at Step 1 or 2 until a solution is found.

Since the cost of correcting defects escalates as the design progresses, the PBD approach can result in poor results by performing the design process in a sequential-only method (Sobek, Ward, & Liker, 1999). The sequential process leads to incorrect work discovered late and challenges in integration. Delay of work is the main issue associated with the process, since major changes must be made once information is transferred to downstream activities (Ward et al., 1995).

Set-Based Design

The theoretical foundation for SBD was established in Allen Ward's MIT PhD thesis in 1989. His work presented a computer compiling program that would assist a mechanical engineer during the design of various systems. Bridging his research on mechanical systems to the broader context of all product development, Ward proposed two product development fundamentals (Ward, 1989):

1. All products should be designed with all viable options in mind.
2. Options should not be eliminated unless there is a logical reason to do so.

Ward's approach results in a gradual narrowing of the system solution space while investigating different design concepts in parallel. Keeping all feasible options in consideration for as long as possible was accomplished by considering groups of mechanical components as "sets," thus leading to the term "Set-Based Design."

In 1995, Ward et al. described a Second Paradox to how TMC executes its business, which included the following generalities: delaying decisions, communicating ambiguously with its suppliers, and pursuing an excessive number of prototypes. This Second Paradox formed the basis for what Ward's research group defined as a culture of SBCE, in which they were able to explain the paradox between seemingly inefficient sub-steps and the efficient overall process by summarizing the SBCE progress into four steps (Ward et al., 1995):

1. The design team considers "sets" of system solutions by defining options of possible sub-system solutions.
2. Possible subsystem design solutions are explored in parallel using analysis, expertise, and experiments.
3. The design team uses the analysis of each subsystem to gradually narrow the sets of system solutions that are possible.
4. Once the design team has found a preferred sub-system solution, the design does not deviate unless absolutely necessary.

Recently, Ghosh and Seering reviewed the previous 20 years of publications relating to SBD principles and characteristics. They qualitatively surmised that organizations performing set-based product development display two principles (Ghosh & Seering, 2014):

1. Considering sets of distinct alternatives concurrently.
2. Delaying convergent decision-making.

How SBD is executed is a unique process. At the beginning of SBD, the conceptual design is organized into separate sub-spaces along the lines of product form or function that align with individual expertise within the design team (Gray, 2011). During this decomposition phase, design teams establish design variables that represent interfaces between sub-spaces. Design teams identify ranges, or sets, for the interfacial design variables based on experts' opinions of what is possible. With interfaces defined that provide a range of possible sub-systems solutions, sub-space design teams are able to independently and concurrently create their own sub-system designs (Sobek et al., 1999).

During this stage of initial design, enough analysis is performed on sub-systems to identify priority sub-system solutions. After preliminary analysis, design teams meet and review sub-space design solutions to identify solutions that have overlapping (shared) design variable ranges. The overlapping regions represent a design space that is feasible for all sub-space design teams (Bernstein, 1998). During these meetings, design teams communicate their preferences for the originally established design variables. Given preferences of other sub-space design teams, the design groups then re-convene and rework designs to incorporate trade-offs and benefits for overlapping feasible design regions. The entire process is gradually repeated with higher fidelity analysis. This process results in eliminating, or not further investigating, regions of the overall design space that are sub-optimal to the whole group (Ward et al., 1995).

An organization that displays both principles can be labeled as utilizing set-based product development. Tailoring the principles to a process for larger complex systems like ship design and acquisition, we offer the following as general principles of SBD:

Principle 1: Establish the design space and sub-divide along areas of expertise: concurrent subsystem evaluation

Principle 2: Gradually and deliberately reduce the design space by integrating preferred subspaces: discovery by elimination

The main benefit of SBD is that it forces teams of designers to communicate in an effective and efficient manner along the lines of product architecture and interfaces—performing design using Principles 1 and 2. SBD communication enacts a decision-making process that enables effective and logical decisions to be made with confidence. Fundamentally, SBD is a design method that discovers the optimal solution by a gradual elimination of the design trade space. The potential benefits of SBD can be summarized as the follows:

- Reduction of later stage rework when the cost of change is more expensive; therefore, less costly to design, build, and maintain the product
- Reduction of design cycle-time; therefore, less costly to design the product and more market share gained from entering an opportunity market sooner
- Better design knowledge capture; therefore, less costly to incorporate customer changes during design or to perform future similar product designs
- A better solution is found because of the methodic reduction of the design trade space; therefore, higher customer satisfaction

Traditional Ship Design

Figure 1 presents a ship design example for a surface cargo ship. This ship design spiral has been ship design tradition since originally presented in 1959 by J. H. Evans.

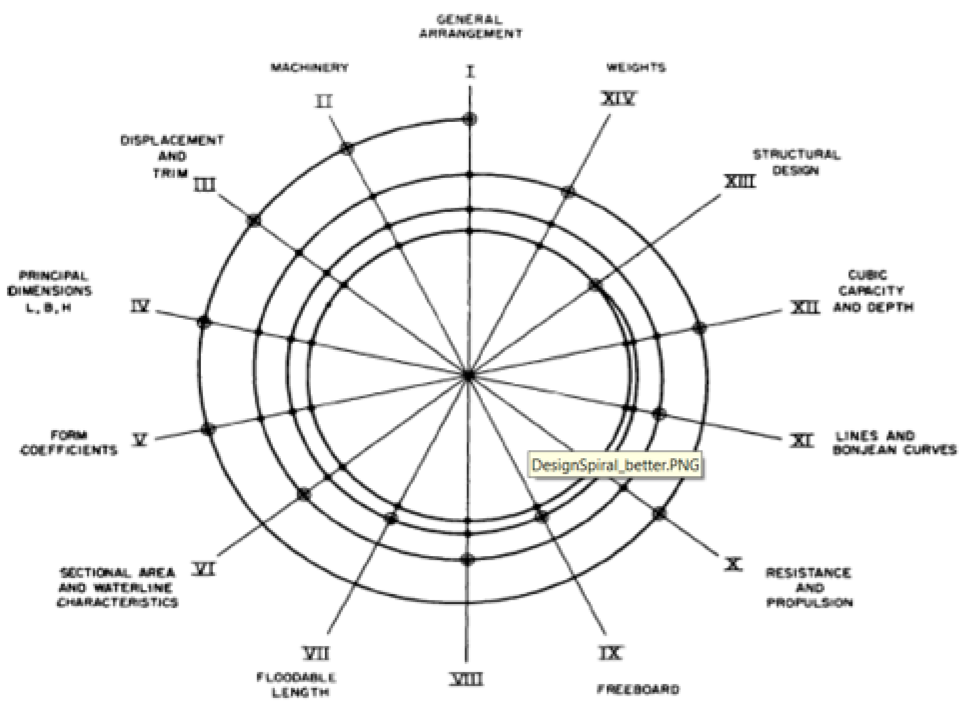


Figure 1. Ship Design Spiral
(Evans, 1959)

This model recognizes the complex nature of the ship design and approaches the design process from the view of conducting iterative passes from one element to the next: weight, volume, stability, resistance, powering, strength, and so on. Systematically addressing each element in sequence, and doing so in increasing detail in each pass around the spiral can reach a single balanced design that satisfies all constraints (Frye, 2010). The model incorporates most of the product development process for ships. Iterations around the spiral would first be performed at the concept design level and gradually proceed toward detailed production design. What aren't captured in Evans's model are the operations and support phases of ship product development.

This approach to ship product development is synonymous with the term *point-based design* since each pass through the spiral attempts to resolve conflicts between elements and develop a design that meets requirements.

Research Methods

Information for this study was gathered by conducting research of open source literature and unclassified databases. Databases accessed were contained on website servers for the Defense Acquisition Management Information Retrieval (DAMIR), Assistant Secretary of the Navy Research, Development, and Acquisition Information System (RDAIS), and USN Visibility and Management of Operating and Support Costs (VAMOSC) systems. In some cases, program-specific documents were classified as Unclassified/For Official Use Only or releasable to only DoD employees or contractors. These documents were only used to identify potential candidates for interviews and are not referenced or cited in this work.

Research data was also obtained through interviews with various stakeholders, decision-makers, sponsors, managers, and engineers within the DoD and Department of the Navy (DoN). Twenty-one interviews were conducted in support of this work from individuals in ASN RDA, OPNAV, CAPE, NAVSEA 05, NSWC-CD, PEOSHIPS, PEOSUBS, CSRA/DoN contractor, and SSGC. Interviewees were asked general questions to understand what the USN values and when in the ship design and acquisition process. More specific questions were then asked to understand processes and tools used to perform respective parts of ship design and acquisition processes.

Case Studies of SBD in the USN

The USN has recently experimented with the use of SBD in acquisition programs. The following examples are programs that actively tried to apply principles of SBD. They are the Pre-PD on Ship to Shore Connector (SCC), Pre-Analysis of Alternatives (AoA) design for the Amphibious Combat Vehicle (ACV), and Pre-AoA design by the Small Surface Combatant Task Force (SSCTF). In each case, some of the author's principles of SBD were identified and some proposed benefits of SBD were achieved.

Ship to Shore Connector (SSC)

The SSC program was created to produce a replacement for the Landing Craft Air Cushioned (LCAC) amphibious transport vehicle. LCACs were designed in the late 1970s and produced during 1984 through 2000. LCACs are still in service today with the oldest LCACs expected to begin retirement in 2019. When considering options for maintaining LCAC amphibious landing capability, the USN performed Exploratory and Pre-AoA design studies in 2006 that resulted in an approved Initial Capabilities Document (ICD) and AoA in 2006 and 2007, respectively (Mebane et al., 2011).

Like other USN ship AoAs, the preferred AoA variant did provide enough detail to satisfy producing the Capabilities Development Document (CDD; Singer, Doerry, & Buckley, 2009). Thus, Pre-Preliminary Design was performed to support refining the draft CDD. NAVSEA ship design leadership decided to pioneer using SBD on the LCAC replacement in accordance with in-progress design process improvement initiatives. These early studies established the LCAC replacement as the SCC program under PMS-377 with a Ship Design Manager (SDM) from NAVSEA 05D. USN leadership was aware of the proposed benefits of SBD, but OPNAV and PMS-377 were most interested in SBD's advantage of critical design decision knowledge capture (McKenney & Singer, 2014) because of the expected high military leadership turnover during typical USN ship design and acquisition (Mebane et al., 2011).

How SBD on SSC Was Executed

Without a formal process described in any USN instruction for SBD, the SCC project team utilized the Decision Object System Engineering (DOSE) method to guide their process for decision-making with the support of experienced academics and consultants familiar with SBD. DOSE's use of knowledge-mapping techniques facilitated team decision-making along lines of functional expertise (Buckley & Stammnitz, 2004; CDI Marine, 2009). With a method to guide overall design execution, the SDM assembled and partitioned the SSC design team per Figure 2 and structured the execution of SSC SBD in three generic phases: (1) Trade space setup and Characterization, (2) Trade space reduction, and (3) Integration and Scoring (Mebane et al., 2011).

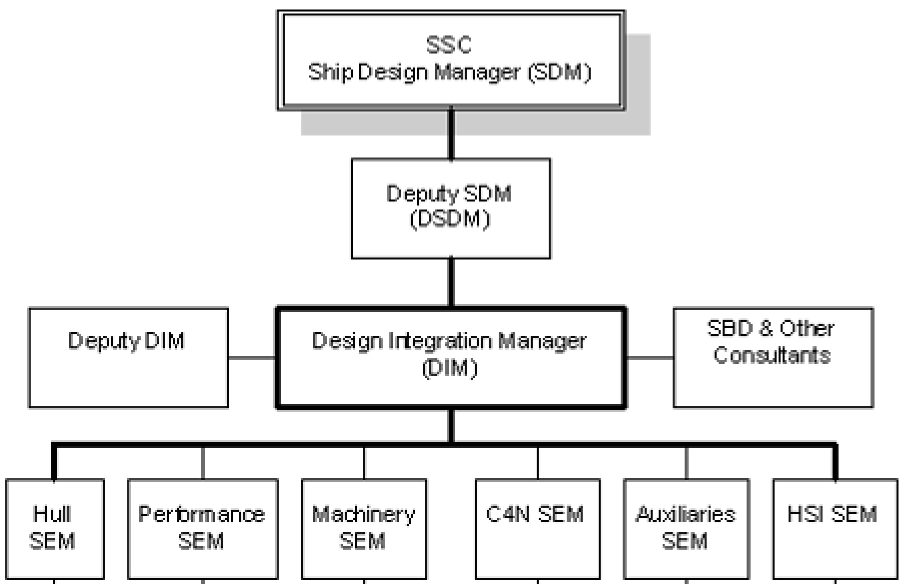


Figure 2. SSC Design Team Structure
(CDI Marine, 2009)

- Ship Design Manager (SDM): The lead system engineer on the project. This individual represents the design team in all matters with outside organizations.
- Design Integration Manager (DIM): This individual is responsible for facilitating communication, decision-making, and integration among all the elements.
- System Engineering Manager (SEM): These individuals represent the system expert in the specific element field.

Trade Space Setup and Characterization

The inputs used for the design effort were shaped into what the design team referred to as a craft-level Functional Design Document (FDD), which was a compilation of NAVSEA executive guidance, the SSC Analysis of Alternatives (AoA) and the SSC Initial Capabilities Document (ICD), and Landing Craft Air Cushion (LCAC) Service Life Extension Program (SLEP) requirements and lessons learned. Using the performance attributes identified in the FDD, Air cushion vehicle Design Synthesis Model (ADSM[1]) was used to convert overall craft performance into performance ranges for each Element: Hull, Machinery, Performance, Combat/Command/Control & Communication Networks (C4N), Auxiliaries, and Human System Integration (HSI). These Element performance ranges were converted into Functional Requirements Documents (FRDs) to guide Element trade space characterization and analysis. When characterizing their trade spaces, SEMs were given latitude to explore any potential solution as long as they had concurrence from a Technical Warrant Holder (TWH) that the proposed system solution was acceptable. At the end of Element characterization, the SEM had a Trade Space Summary (TSS), in the form of an MS Excel spreadsheet, which captured TWH comments, approvals, and future trade space reduction decisions.

Trade Space Reduction

After establishing Element solution trade space acceptability, SEMs used design of experiments, or other analysis, to analyze their intra-element set of solutions for key design parameter preference or dominance. Model Based System Engineering techniques compared intra-element solutions against each other by identifying performance measures, modeling and simulation scenarios appropriate for each element based on Subject Matter Expert (SME) opinion. Some SEMs used Response Surface Methodology to compare alternatives, where others used a less rigorous approach because of the lack of design variable continuity over the FRD. This process was completely concurrent for each SEM and was supervised and facilitated by periodic Design Integration Team (DIT[2]) meetings. TSSs captured these reduction decisions. At the end of the trade space reduction phase, each SEM had a set of non-dominated intra-element solutions. These solutions were approved by TWH's as technically acceptable and concurred upon by the DIT as viable. The next step in the SBD design effort was to combine all Element solutions into craft variants.

Integration and Scoring

Towards the end of Trade Space Reduction, the DIT identified what they referred to as "negotiating relationships" between Elements, which resulted when the selection of one option in an Element influenced which options could work in the other Elements. Eliminating exclusions based on negotiating relationships resulted in the set of all potentially viable SSC crafts. Next, all potentially viable craft designs were submitted to a Balancing Process in which a design synthesis tool, similar to ADSM, was performed for each candidate craft to ensure design candidates pass a first order test for platform viability. For the SSC project, the balance process screened candidates for important high-level craft attributes: an initial stability check, a test for adequate power to get over the generated bow wave, and a test for

---

[1] ADSM is an air-cushioned craft-specific design tool created by TMLS and maintained by the USN for LCAC/SCC design

[2] The DIT consisted of the DIM, the Deputy DIM, and SBD consultants.

adequate power to maintain the required cruise speed. The balancing process eliminated another significant portion of SSC alternatives and produced a set of metrics for each variant that could be used for quantitative comparison. A scoring scheme using an Overall Measure of Effectiveness (OMOE) from multi-attribute utility was created to evaluate the remaining SSC variants between cost, risk, and performance. This resulted in a small group of high scoring variants in which the design team chose two variants, which only differed by hull material selection, to carry into Preliminary Design. Figure 3 captures the three phases of the SSC design process.

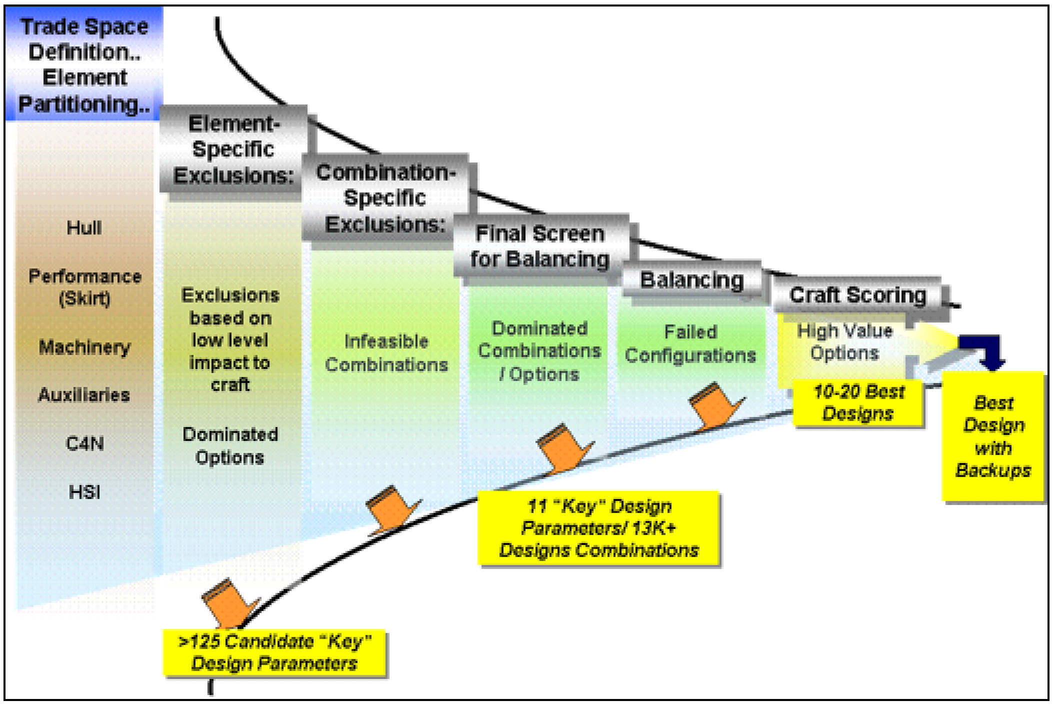


Figure 3. SSC Set Reduction Process
(CDI Marine, 2009)

Results of SCC SBD

At the end of design, two preferred, similar variants were identified by the team as the believed global optimums (Mebane et al., 2011; Singer et al., 2009). Additionally, a vast amount of design knowledge had been captured using the TSSs, specifically the negotiating relationships. The two SSC variants were generated by a process that evaluated functional specific trade spaces concurrently and reduced the trade space by eliminating dominated or infeasible options; thus, satisfying the authors' principles for SBD execution.

Programmatically, the SSC design was completed on time, within 10% of budget, and used little to no design margins (Doerry, 2010). Overall, in 2008, the SBD results for SSC were immediately used for Preliminary Design (PD), Contract Design (CD) and Gates 4/5 of the newly instituted two pass/six gate (2P/6G) process. The SSC program has proceeded past MS B and is supervising construction of the SSC test craft at Textron Marine and Land Systems.

Other Benefits of Using SBD in the SSC Program

Although the overall process used by the SSC team may not have been textbook SBD (McKenney & Singer, 2014), the USN was writing the textbook for using SBD during

SSC design efforts (Singer et al., 2009). The biggest lesson learned from the SSC Pre-PD design phase was that SBD principles could be translated into a process for use on USN ships/crafts. The SBD process not only quickly (within four months) produced results, but also the formal decision-making exclusions and eliminations provided excellent design knowledge capture. For SSC, knowledge capture was obtained by TSSs and the eventual discovery of negotiating relationships between functional Element groups. This design knowledge capture also led to a more fluid design review process during PD, CD, Gate 5, and MS B. The design team was able to immediately answer, or in even some cases prevent, design reviewer and higher level decision-maker questions about the recommended edges of the SCC design. Once design reviewers and higher level decision-makers understood the trade space elimination process, they became satisfied that an ideal solution had been reached. Thus, there was no need to further question why the design team arrived at their recommended solution. In the end, future fluidity of design review is what the USN hopes to achieve by capturing lessons learned from the SSC.

Design fluidity, in the SSC case, could be translated into better cost and schedule performance. By preventing the extra questions from reviewers during the PD, CD, and Gate 5 review phases, the SSC design team ultimately prevented undertaking additional studies to answer posed questions. In the past, these extra questions were considered to be significant due to either the seniority level that asked the question or because the question was generated in front of a large diverse group. After performing the study to answer the extra question, the conclusion often ended up being low value re-work.

Amphibious Combat Vehicle (ACV)

Another opportunity arose for a rapid ship/craft design learning event shortly after the SSC team finished their critical design review. In 2011, the United States Marine Corp (USMC) canceled the 40-year old Amphibious Assault Vehicle (AAV) replacement program, the Expeditionary Fighting Vehicle (EFV), due to poor reliability and excessive cost (O'Rourke, 2016a). The USMC immediately began re-planning for the development of a more affordable and sustainable amphibious combat vehicle (ACV). This resulted in an ICD to align capabilities and future CONOPs and an AoA that re-enforced the need for a self-deploying survivable craft. But neither the ICD nor AoA explored the operational benefits of a high water speed (HWS) craft (Burrow et al., 2013). With extra scrutiny on the ACV program from the previous EFV cancellation, senior USMC leaders expressed concern with proceeding with a low water speed craft without evaluating the HWS requirement, citing operational flexibility and the potential tactical advantage HWS might have (Burrow et al., 2013). To satisfy the "what about…/what's not shown on the slide" question from USMC leadership, the Assistant Commandant of the Marine Corps and ASN RDA developed an ACV directorate team to evaluate the cost and capability trade-offs of a HWS ACV.

The ACV design team was focused on expeditiously answering the "what about …" question while simultaneously using previous EFV information and capturing ACV design knowledge. The proposed reduced cycle-time and knowledge-capture benefits of SBD aligned with the ACV directorate's priorities. Therefore, the ACV design team desired to explore incorporating aspects of SBD, where possible, in the ACV design approach. In the end, the results of the ACV design produced a detailed cost–benefit assessment of the HWS requirement. Additionally, USN and USMC leadership became more aware of configuration diversity terminology and how early stage design decision information may be presented using a SBD approach.

How the ACV Design Was Executed

To assess the feasibility and cost of the HWS ACV, the ACV directorate established a design team, formulated an analysis plan, and executed a series of four focused design studies. Where possible, concurrent design efforts were performed, and design knowledge was shared between core teams to improve the validity and value of sequential ACV design studies.

With clear performance requirements, the ACV team generated a library of ACV components that could comprise an ACV variant based on the AAV work breakdown structure. The library initially incorporated only proven low-risk technologies, but was expanded to high risk/high reward components based on the Innovation Team's research (Burrow et al., 2013). Component size, weight, and cost information was the basis for the Market Research Database (MRDB), which utilized the synthesis tool Framework for Assessing Cost and Technology (FACT) for evaluating ACV performance. With a library of components, a set of requirements, operational scenarios, and a performance synthesis model, the ACV design team was able to generate a large design trade space of potential configurations to satisfy capability concepts.

To evaluate the large trade space, individual studies were performed to first validate the design team's models and then to target specific design attributes. The Baseline Study was performed to validate the process models and design tools. The follow-on studies further explored technical viability and specific operational performance of HWS vs LWS ACVs. The four studies used multi-attribute utility theory to produce configuration Performance vs. Cost graphs.

The ACV design team claimed their use of diversity in design decisions was set-based; but the use of the SBD principles described in the section on the ACV was sparse. The only aspect of the SBD principles that occurred during the ACV design was the knowledge-sharing that occurred between the functional groups. This partial use of SBD has been identified by some researchers as aligning with effective trade space exploration (Ghosh & Seering, 2014; Schmid, 2015) and is a better description of the overall design approach used by the ACV design team. As the requirements group identified new or changing requirements from the USMC, they would update DOORS. A DOORS update changed the parameters of FACT, which then ultimately resulted in opening or eliminating some of the ACV configuration trade space. Additionally, as the Affordability Analysis team identified supply chain or logistic issues that resulted in the preference of one component over the other, the MRDB would be updated. Changed parameters in the MRDB resulted in configuration utility changes, which could impact final recommendation results. This knowledge-sharing represented separate groups of concurrently evaluating sub-systems (Principle 1). Outside of the SBD principles described previously, the ACV directorate introduced the topic of cost diversity in which the overarching SBD premise of the optimal solution residing within the feasible set was reinforced.

What Was Learned From ACV Design

Overall, the ACV design assessed HWS ACV feasibility and cost. The design team felt they achieved this goal by performing design in a way that produced presentable, understandable information to decision-makers. They felt the presentation of design information supported a high degree of confidence in cost and risk decisions. Interviews with ASN RDA and reviews of the literature confirmed what the ACV team believed, that leadership was very satisfied with the ACV design team results (ASN-RDA, 2016). In the end, the ACV team was able to address leadership "what if" questions succinctly and with the technical rigor to enable high confidence decisions. Additionally, the ACV concept design introduced and familiarized USN leadership with a design information presentation style founded on solution feasibility, viability, and diversity discovered through a SBD approach.

Ultimately, the USMC selected the LWS ACV configuration as the initial, affordable selection as part of an incremental acquisition strategy that could eventually include a HWS variant.

Small Surface Combatant Task Force (SSCTF)

On February 24, 2014, Secretary of Defense Chuck Hagel restructured the LCS program by directing the USN to provide alternate proposals to procure a more capable and lethal small surface combatant for the last 20 of 52 planned LCSs (O'Rourke, 2014). Originally, the LCS program was announced in 2001 as a variant of the Future Destroyer concept of operations amid the large decisions facing the USN after the 1997 Quadrennial Defense Review (O'Rourke, 2016b; Work, 2007).

In the spring and summer of 2014, the USN responded to SECDEF's LCS restructure direction by assembling a group of surface warfare, ship design, and industry experts: the Small Surface Combatant Task Force (SSCTF). The SSCTF received direction from ASN RDA and the CNO to (Garner et al., 2015):

- Establish the requirements for a small surface combatant
- Assess the requirements delta against the existing LCS (both sea frames)
- Translate the requirements delta into concept designs considering: existing ship, a modified existing ship, and new ship design options with schedule, cost, sensor systems, and lethality measures of performance.

Similar to the ACV concept design, one of the priorities for USN leadership was quickly coming to a well-informed decision to re-direct a program proceeding in the wrong direction. Fresh from the ACV concept design experience, a core group of NAVSEA 05D SDMs were available to advise the SSCTF on use of SBD in concept design. Their insights enabled the SSCTF to tailor their design approach to take advantage of the knowledge-sharing and concurrent work principles of SBD. The overall approach the SSCTF used to achieve their tasking: (1) capabilities were defined, (2) capabilities were translated into configurations of different ship systems to achieve required capability performance levels, and (3) synthesized ships were evaluated using utility theory for performance vs cost (Garner et al., 2015). During this effort, the SSCTF utilized the SBD principles described previously during synthesis and evaluation.

How SBD Was Used During SSCTF

One of the SBD principles used by the SSCTF design team was concurrent design of the Hull, Mechanical and Electrical (HM&E) and Combat Systems during synthesis. When designing and converging full ship designs, HM&E experts assumed the Space, Weight, Power and Cooling (SWAP-C) metrics for the combat system. HM&E designers utilized a large (low-risk) range for combat system SWAP-C architecture to more likely enable future ship convergence feasibility and therefore viability. Establishing these "placeholders" for combat system architecture allowed the combat warfare system experts to independently design their systems. As combat system design solutions matured, the matured combat system SWAP-C metrics were intersected with the HM&E assumptions to refine the solution space. Performing the HM&E and combat system work in parallel and then intersecting design efforts matches the first and second principles of SBD from the Set-Based Design section. The SSCTF claimed to follow the literary SBD principle of canvassing a large trade space, but just because a large trade space is generated at the onset of design doesn't make a design approach set-based.

What Was Learned From SSCTF Design

Three major points were learned from the SSCTF design effort. First, early stage ship SBD can be achieved by partitioning along HM&E and Combat Systems functional boundaries. The interfacial variables that exist between these two groups are physics-based variables which are easily quantified within existing design tools. Furthermore, USN ship design tools such as Advanced Surface Ship and Submarine Evaluation Tool (ASSET), Rapid Ship Design Environment (RSDE), and Leading Edge Architecture for Prototyping Systems (LEAPS) provide effective, rapid generation and comparison of ship designs independent of concurrently working in the HM&E or Combat System functional group. These tools easily intersect interfacial variables between the HM&E and Combat System functional groups. Second, USN leadership preferred the visual risk assessment and data presentation that accompanied the ASSET, RSDE, and LEAPs design products. Similar to the ACV design, USN leadership discussed their perceived confidence in decision-making based on the in-depth and easily decipherable data presented by SSCTF designers. Third, the USN ship design community has established a core group of designers that can responsively react to emergent ship design tasking and produce well received results in a rapid fashion. Overall, as the most recent SBD ship excursion, the SSCTF has helped validate the tools, processes, and metrics associated with a set-based surface ship design.

Summary of USN Cases of SBD

The four proposed benefits of SBD identified previously are assessed using available programmatic information for SSC, ACV, or SSCTF to determine whether evidence supports the claimed benefit. The proposed benefits of SBD are as follows:

- *Reduction of later stage rework when the cost of change is more expensive; therefore, less cost to design, build, and maintain the product*. This benefit did not specifically appear in literature or interviews for the SSC, ACV, or SSCTF, but can be inferred indirectly. Each of the USN SBD cases occurred early[3] during the ship product development life-cycle. Therefore, the overall acquisition cost performance of the ship program should be improved based on the SBD principle of reducing later more costly re-work. This can be assessed by reviewing the adherence of a program's actual acquisition cost to its original APB cost in a SAR. The ACV and modified-LCS have not proceeded past their MS B APB decision, so only the SSC can be assessed. Cost performance is captured in Figure 4 and shows that the SSC has the highest acquisition cost performance; achieving greater than 1.0 means that overall actual acquisition costs have decreased compared with the original APB estimates.

[3] Even though the SSCTF design event occurred during mid-life of the LCS, it was evaluating ship design concepts from the beginning of the ship product development life-cycle.

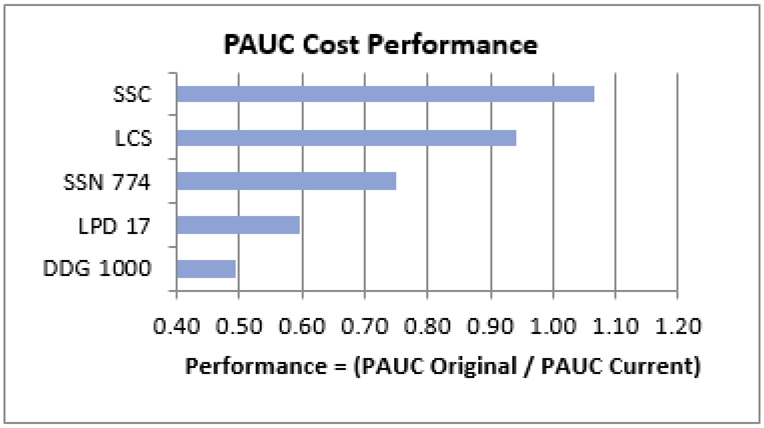


Figure 4. Per Unit Acquisition Cost Performance
(DAMIR, 2015)

- *Reduction of design cycle-time; therefore, less cost to design the product and more market share gained from entering an opportunity market sooner*. For each of the USN design cases studied, design cycle-time reduction was a goal of each design team. All three design teams reported producing results within a time-span previously not achievable. USN leadership confirmed the previously not achievable claim for each case. Therefore, this proposed benefit was realized by USN's use of SBD.
- *Better design knowledge capture; therefore, less costly to incorporate customer changes during design or to perform future similar product designs.* Knowledge capture was also a goal of each design case studied. In the SSC design, this attribute was realized by generating TSSs and the identification of negotiating relationships. The design team specifically stated that future modernization or recapitalization efforts for SSC would go smoother with the information gathered during SSC design. The likelihood of knowledge capture improving future design or cost performance for the ACV and SSCTF still remain to be seen. In ACV and SSCTF designs, the SBD discovery by elimination principle was not specifically adhered to throughout each design. Therefore, only time will tell if the USN will fully realize this proposed SBD benefit. This proposed benefit also hints at SBD being a more flexible design approach due to the ability to easily re-open parts of the design space that were previously excluded by a elimination/design space reduction decision. Although changing stakeholder or customer requirements was not highlighted by literature or reported as significant during interviews for each USN SBD case, this topic was discussed during a general interview with NSWC CD. NSWC CD reported conducting an internal study in which two different teams independently used SBD and PBD to design a surface combatant. The results of their study showed that the PBD team needed significantly more rework to accommodate requirements changes and the mid-life upgrade modernization (Gray, Rigterink, & McCauley, 2017). Therefore, this proposed benefit has been demonstrated in a structured ship design academic setting.
- *A better solution is found because of the methodic reduction of the design trade space; therefore, higher customer satisfaction*. To evaluate this benefit, the "customer" needs to be defined. For the three design cases, the customer could be acquisition leadership, or the USN sailor who will eventually operate

the warship product. Additionally, it is difficult to gauge the degree of satisfaction in both the acquisition leadership and USN sailor cases. As best determined from interviews with executives involved with the three cases studied, leadership was satisfied with the results of each design. They reported being able to better understand information presented during design update or final briefs based on the visual representation that accompanied the Monte Carlo simulations for ACV and SSCTF and the key variable reduction graph from SSC. Data presented in this manner was able to drive home the SBD principle point about designing through discovery by elimination.

Overall the case studies of USN SBD use showed that at least some of the SBD principles were adhered to. A major learning is that USN leadership has become more familiar and accustomed to the style and depth of early stage design information resulting from a set-based design approach. The proposed benefits of SBD were shown to have been realized in some capacity for each example. Last, SBD is acknowledged in USN ship design process instructions and SBD ship design tools continue to be developed in the USN.

New Path Forward Using SBD in Ship Design

The Secretary of the Navy has authored and issued SECNAV 5000.2E to depict how the USN will operate within the DAS/PPBE/JCIDS triad and describe the 2P/6G process for ship product development. The purpose of the 2P/6G process is to improve insight into ship development and execution of its acquisition (NAVSEA, 2010). For each gate, stakeholders and their priorities were identified through interviews with key stakeholders in the process, summarized in Table 1. Reflection on the priorities and outcomes of each gate suggest the following:

- SBD is not appropriate for Gate 1. The principles of SBD do not align with desired outcomes of the CBA and ICD because the CBA focuses on using existing ship designs and plans, and the conduct of an ICD doesn't require the sophistication or detail generated by SBD.
- The principles of SBD align with stakeholder priorities for Gate 2. The SBD principle of concurrent sub-subsystem evaluation aligns with the desire for trade studies.
- SBD is potentially the best method of providing what stakeholders want from Gate 3. Understanding the "drivers" of KPP/KSA cost requires that the ship designer understand "relationships" between systems that cause weight, which is what SBD does inherently.
- SBD may be amenable to activities during Gate 4 depending on the design progresses used in Gate 2 to Gate 3. Because the ship design is already partitioned and close to complete, utilizing SBD may or may not provide additional benefits over point-based or traditional ship design. The design method used during Gate 3 activities is likely to be the best to use during Gate 4.

Table 1. Gate 1–4 Stakeholder Priorities

Gate	What	Type of Design	Who	Priorities
1	ICD	Exploratory & Pre-AoA	N8	Understand intelligence & technology risks. Don't jump to conclusions.
2	AoA	AoA	CNO	Large span of AoA variants. AoA cost vs. capability trade-off information.
3	CDD	Pre-PD	CNO	Feasibility assessment of KPP/KSA performance values.
4	SDS	Preliminary	ASN RDA	Cost and feasibility of sub-system integration.

Gate 2 SBD Process Improvement: The Analysis of Feasibility

During interviews, discussions, and literature research, the AoA seemed to be the first large decision point and possibly the most influential in directing the course of a ship design. For ship design programs, the AoA occurs right at the time of the highest design influence on cost and capability. Almost every ship program researched had a foundational AoA that, at the very least, collected and described the basic need for the ship. Given the importance of the AoA in the ship design, it was surprising to discover that past USN ship AoAs have presented a limited and scripted selection of options that often fail to carry forward as the program progresses. This is because they are generally sparse point designs that offer neither a useful range of options nor useful insights to migrate toward a more optimal design. Given the amenability of SBD to Gate 2 stakeholder priorities and the historical poor performance of USN ship AoA's, the current ship AoA process was assessed for SBD process improvement opportunities.

Current ship design tools can support a different way to perform a ship AoA. This new approach uses RSDE's capability to communicate interfacial design variables to achieve the principles of SBD. This new method, termed the Analysis of Feasibility (AoF), improves the current AoA process by producing data that better aligns with DoD 5000.02 AoA guidance, eliminating the "middle point" pitfalls of past AoAs, and providing results in response surfaces instead of bar charts. Additionally, the AoF enables follow-on pre-preliminary design to continue in a set-based fashion. Most importantly, the AoF produces results in a fashion preferred by stakeholders to enable higher confidence decisions. Lastly, the AoF contributes to lower overall PAUC by preventing future re-work and shortening overall ship design cycle-time.

A process, similar to that used by the SSCTF, can be implemented using existing ship design tools and a set-based approach to improve Gate 2 activities. The SSCTF used two functional teams split between Combat Systems and HM&E to accomplish capability concept designs. The same AoF design team division could be used by a NAVSEA SCM, who is familiar with SBD, to generate a large span of variants to inform a ship AoF trade-off study.

Ship AoA's use the Design Reference Mission (DRM) from the ICD to determine required ship performance. For ships, the DRM determines the type and variety of Combat Systems, but not the sea frame that carries it. A simple analogy is to think of the Navy ship as a truck which carries the sensors and weapon systems to perform the DRM. The truck supplies the weapon systems with energy and physical support. Splitting a design team along the weapons system and truck functional boundaries would establish energy and physical support as interfacial variables; and thus, partition the design space into separate groups of experts (SBD Principle 1). By splitting into two functional teams, Ship and Combat System designers can independently and concurrently design sets of systems that meet the required performance of the DRM. Once design sets are complete, the two teams meet/communicate to share what range of energy and support each team needs from the other (SBD Principle 2). For example, determining how the truck is built determines how heavy or high the weapons system could be placed before the truck tips over or breaks. Likewise, the DRM would determine the size and type of weapons system needed.

In the USN ship design environment, ASSET has the capability to design the ship [truck] and place the weapons system. Bu, ASSET produces results for only one unique, individual ship and weapons system configuration at a time. Using ASSET with RSDE allows a range of ship design parameters and a range of weapon system locations and sizes to be analyzed concurrently. Figure 5 shows a visual description of how a range of combat system configurations could be varied while simultaneously varying ship parameters.

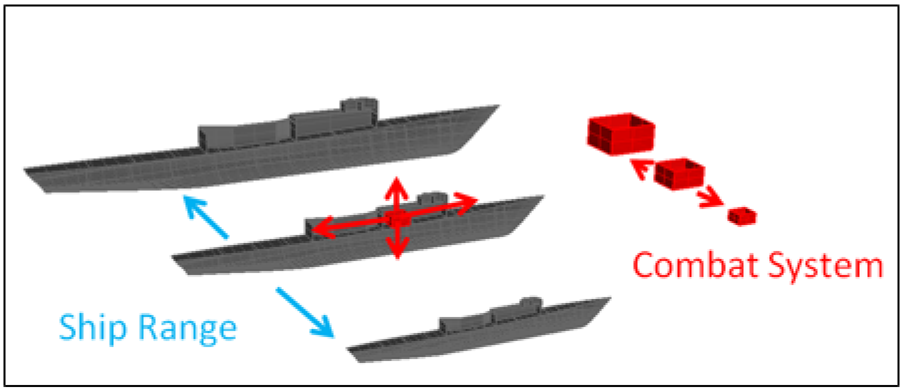


Figure 5. New AoF Variant Creation Process

Similarly, various different propulsion, power generation, and auxiliary cooling configurations are typically considered when performing early stage ship design when all available options should be included in the design trade space. Each set of ship parameters can be evaluated over the range of possible Combat Systems by simply utilizing RSDE to run multiple ASSET evaluations

Overall, conducting an AoF in a set-based manner produces a large number of variants that would fill in the middle points of a current state AoA. This larger data set should produce better capability and cost trade-off assessment for decision-makers using statistical tools like JMP. JMP can easily produce graphs and view charts that quickly show regions of the performance variables and how they change with variations in Ship or Combat System parameters. Most importantly, data presented in this manner is preferred by stakeholders over the classic cost and capability bar-charts. The concurrent evaluation of the trade space by the Ship and Combat System design experts should result in a faster design cycle-time. Also, the knowledge obtained by identifying the ranges of infeasibility for various Ship parameters and Combat System configurations is invaluable. This design knowledge is captured by the formal meeting/communicating process inherent to SBD and supports eliminating future more-costly re-work. Overall, using a set-based AoF approach in Gate 2 should support a lower ship program PAUC and faster acquisition.

Gate 3 SBD Process Improvement: Continue the AoF Analysis

Given the amenability of SBD to the stakeholder priorities of Gate 3, the AoF method provides an opportunity to improve Gate 3 ship design activities. With the priority of assessing the feasibility of KPPs/KSAs, the large trade space and design knowledge gained from Gate 2 AoF activities presents an excellent opportunity to support continued design feasibility assessments.

After a successful Gate 2 review board, the focus of the ship design team turns towards generating the CDD. Which ultimately means conducting enough design to assess if the performance levels required in the KPPs/KSAs can be achieved within cost and schedule targets. In the past, AoAs have produced preferred variants that do not represent the right combination of affordable capability; thus, ship program sponsors have had to fund re-design efforts on AoA resultant variants. These redesign efforts present an improvement opportunity for Gate 3. Continuing the AoF design method can reduce or prevent the re-design effort experienced in past Pre-PD and PD designs.

The AoF design method reduces or prevents AoA variant re-design by keeping the design trade space open across the Gate 2 review board. In the past, AoAs were contained design events that only produced a written report to make a decision. AoA ASSET ship models were retained by NAVSEA 05D, but rarely re-used because exact AoA variants tended to not exactly align with what the ship program manager desired for the CDD.

To continue the AoF in Gate 3, the SDM would start by re-evaluating the design team functional partition to identify sub-regions of expertise for further concurrent evaluation. For example, in the Gate 2 AoF, the propulsion system was only at the “type” level. A specific Propulsion design team could be created during Gate 3 with identified interfacial design variables of space, thrust, and weight with the Ship design team. This would support evaluating “options” of different propulsion methods. Once a span of propulsion options has been studied, the propulsion team would communicate the exact space, thrust, and weight of their preferred propulsion choice to the Ship team. This would most likely eliminate some of the propulsion options from consideration and thus refine the performance of the overall ship. The ship design tools ASSET and RSDE could be used to perform this type of sub-group study. Figure 6 highlights the Propulsion design team example communicating across the identified interfacial variables.

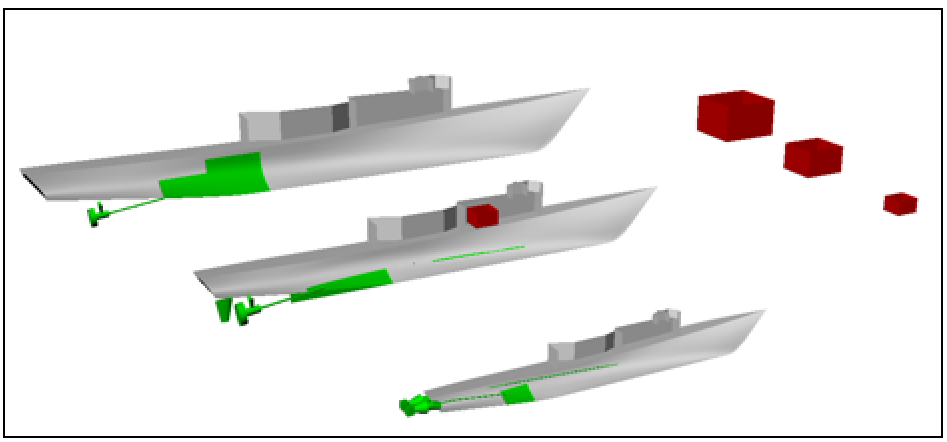


Figure 6. Propulsion Option Exploration

Overall, continuing the AoF approach in Gate 3 provides the SDM with the opportunity to flexibly adjust the design team in areas of the design that may need more specific evaluation to provide a feasible assessment of required KPP/KSA performance. Also, it limits the iterations of re-design performed in the past by keeping the real design trade space open and using SBD to reduce or eliminate dominated sub-system options. Thus, the AoF represents a potentially better way to proceed through Gate 3 activities.

Conclusion

This discussion illustrated how the stakeholder priorities of Gates 1–4 can be addressed by SBD. Gates 2 and 3 were identified as the most likely candidates for SBD process improvement. The Analysis of Feasibility, using SBD principles and existing modern ship design tools, was introduced as a way to improve overall ship program PAUC and design cycle-time by segmenting the to-be-designed ship initially into Ship and Combat System functional design teams. The AoF method illustrates the capability to keep the ship design space open across the Gate 2/3 boundary. In sum, the AoF method uses existing ship design tools and SBD principles to deliver Gate 2 and 3 stakeholder priorities in a preferred fashion.

Conclusion

The proposed process improvement initiatives described in this work are within the capability of the current USN ship design and acquisition workforce. Future work might entail the development of new written policy and guidance at an institutional level. Furthermore, the USN should continue its investment in ship design and process tools that align with the principles of SBD. The proposed benefits of SBD, as applied to USN ship design, are potentially significant. In the face of near- to mid-term ship acquisition challenges, aligning the amenable aspects of the 2P/6G USN ship design process with SBD is one of the more promising opportunities to realize ship design and acquisition improvement.

References

ASN-RDA. (2016, January 22). Executive.

Bernstein, J. I. (1998). *Design methods in the aerospace industry: Looking for evidence of set-based practices* (Master's thesis). Cambridge, MA: Massachusetts Institute of Technology.

Buckley, M., & Stammnitz, V. (2004). Decision oriented systems engineering (DOSE)—A Structured, systematic approach to function allocation. *14th Annual International Symposium for INCOSE.*

Burrow, J., Doerry, N., Earnesty, M., Was, J., Myers, J., Banko, J., . . . Tafolla, T. (2013). Concept exploration of the Amphibious Combat Vehicle. *SNAME Maritime Convention*. Houston, TX. Retrieved from http://www.doerry.org/norbert/papers/20140726ConceptExplorationoftheAmphibiousCombatVehicle.pdf

CDI Marine, B. L. (2009). *Ship-to-shore connector implementation of set based design* (Working Paper 765.002.023-1).

Chapman, W., Bahill, T., & Wymore, A. (1992). *Engineering model and design*. Boca Raton, FL: CRC Press.

DAMIR. (2015, December 15). *Defense Acquisition Management and Information Retrieval.* Retrieved from Selected Acquisition Reports (various programs): http://www.acq.osd.mil/damir/

Doerry, N. (2010). Transitioning technology to naval ships. Washington, DC: NAVSEA 05T.

Evans, J. (1959). Basic design concepts. *American Society of Naval Engineers Journal, 71*(4), 672–678.

Fireman, H., Nutting, M., Rivers, T., & Carlile, G. (1998). LPD 17 on the shipbuilding frontier: Integrated product and process development. In *Proceedings of the Association of Scientists and Engineers 35th Annual Technical Symposium.*

Frye, M. (2010). *Applying set based methodology in submarine concept design* (Master's thesis). Cambridge, MA: Massachusetts Institute of Technology.

Garner, M., Doerry, N., MacKenna, A., Pearce, F., Bassler, C., Hannapel, S., & McCauley, P. (2015). *Concept exploration methods for the Small Surface Combatant.* Retrieved from http://www.doerry.org/norbert/papers/20150717ssc-ce.pdf.

Ghosh, S., & Seering, W. (2014). Set-based thinking in the engineering design community and beyond. In *Proceedings of the ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference,* (pp. DETC2014–35597). Buffalo, NY.

Gray, A. (2011). *Enhancement of set-based design practices via introduction of uncertainty the use of interval type-2 modeling and general type-2 fuzzy logic agent based methods* (Doctoral dissertation). Ann Arbor, MI: University of Michigan.

Gray, A., Rigterink, D., & McCauley, P. (2017). *Point-based versus set-based design method for robust ship design.* Carderock, MD: Naval Surface Warfare Center Carderock Division.

Kassel, B., Cooper, S., & Mackenna, A. (2010). Rebuilding the NAVSEA early stage ship design environment. *American Society of Naval Engineers,* 1–12.

Keane, R., Firemann, H., Hough, J., Helgerson, D., & Whitcomb, C. (2009). Ready to design a naval ship? Prove it! *Journal of Ship Production*, 45–58.

Keane, R., McIntire, J., Fireman, H., & Maher, D. (2009). The LPD 17 ship design: Leading a sea change toward collaborative product development. *Naval Engineers Journal, 121*(2), 15–61.

Liker, J., Sobek, D., Ward, A., & Cristiano, J. (1996). Involving suppliers in product development in the United States and Japan: Evidence for set-based concurrent engineering. *IEEE Transactions on Engineering Management, 43*(2), 165–178.

McKenney, T., & Singer, D. (2014, July). Set-based design: A concurrent engineering approach with particular application to complex marine products. *Marine Technology*, 51–55.

Mebane, W., Carlson, C., Dowd, C., Singer, D., & Buckley, M. (2011). Set-based design and the ship to shore connector. ASNE Day (February 9–10, 2011). Arlington, VA.

NAVSEA. (2010). *Acquisition strategy guide*. Washington, DC: Naval Sea Systems Command.

NAVSEA. (2012). *Ship design manager's and ship integration manager's manual.* Washington, DC: Naval Sea Systems Command.

O'Rourke, R. (2014). Na*vy Littoral Combat Ship (LCS) program: Background and issues for Congress*. Washington, DC: Congressional Research Service.

O'Rourke, R. (2016a). *Marine Corps Amphibious Combat Vehicle (ACV) and Marine Personnel Carrier (MPC): Background and issues for Congress*. Washington, DC: Congressional Research Service.

O'Rourke, R. (2016b). *Navy Littoral Combat Ship (LCS)/Frigate program: Background and issues for Congress*. Washington, DC: Congressional Research Service.

Raudberget, D. (2010). Practical applications of set-based concurrent engineering in industry. *Journal of Mechanical Engineering, 56*(11), 685–695.

Rhodes, D., & Ross, A. (2016, March 4). *Engineered resilient systems—Systems engineering: Knowledge capture and transfer*. Retrieved from Stevens Institute of Technology website: http://www.sercuarc.org/wp-content/uploads/2014/05/SERC-2014-TR045-1EngineeredResilientSystemsRT103-0829FINALmm.pdf

Schmid, R. (2015). *Performance measuring of set-based concurrent engineering in comparison to point-based concurrent engineering using system dynamics* (Master's thesis). Munich, Germany: Technical University of Munich.

Singer, D., Doerry, N., & Buckley, M. (2009). What is set-based design? *Naval Engineers Journal, 121*(4), 31–43.

Sobek, D., Ward, A., & Liker, J. (1999). Toyota's principles of set-based concurrent engineering. *Sloan Management Review, 40*(2), 67–83.

Sullivan, P. E. (2008). *Ship design and analysis tool goals* (Ref: COMNAVSEA memo 9000 Ser 05D/047).

Ward, A. C. (1989). *A theory of quantitative inference for artifact sets, applied to a mechanical design compiler* (Doctoral dissertation). Cambridge, MA: Massachusetts Institute of Technology.

Ward, A., Liker, J., Cristiano, J., & Sobek, D. (1995, Spring). The second Toyota paradox: How delaying decisions can make better cars faster. *Sloan Management Review, 36*(3), 43–61.

Work, R. (2007). *Know when to hold 'em, know when to fold 'em: A new transformation plan for the Navy's surface battle line*. Washington, DC: Center for Strategic and Budgetary Assessments.

THIS PAGE INTENTIONALLY LEFT BLANK

Panel 16. Small Business in Defense Acquisition

Thursday, April 27, 2017	
1:45 p.m. – 3:15 p.m.	**Chair: Emily Harman,** Director, Office of Small Business Programs, Department of the Navy ***Examining the Effects of Set Aside Policies on Competition and Growth for Small and Mid-Sized Suppliers*** Trevor L. Brown, The Ohio State University Amanda M. Girth, The Ohio State University ***Evaluating the Impact of Small Business Set-Asides on Acquisitions Efficiency*** William Lucyshyn, University of Maryland John Rigilano, University of Maryland ***Exploring Drivers of Better Strategic Sourcing in the Air Force Using Analytics*** Aruna U. Apte—Naval Postgraduate School Karen A. F. Landale—Naval Postgraduate School Rene G. Rendon—Naval Postgraduate School Javier Salmerón—Naval Postgraduate School

Emily Harman—is the Director, Office of Small Business Programs (OSBP), for the Department of the Navy (DoN) serving as chief advisor to the Secretary on all small business matters. She is responsible for small business acquisition policy and strategic initiatives.

Harman joined the Secretary of the Navy Staff as a member of the Senior Executive Service in August 2015 and has over 30 years of federal service. Prior to receiving this appointment, she served as Associate Director of the Naval Aviation Systems Command's (NAVAIR's) OSBP from November 2005 to August 2015.

Harman's previous experience includes serving as a Division Director in the Major Weapons System for Air-Antisubmarine Warfare, Assault, Special Mission Programs Contracts Department, and as the Multi-Mission Helicopters Program Office's (PMA-299) Contracting Officer. Harman has NAVAIR experience as a Services Contracting Officer, as well as Contracting Officer for the AV-8B Weapon Systems Program Office (PMA-257).

Prior to joining NAVAIR in 1997, Harman served as a Contracting Officer for the Naval Supply Systems Command's (NAVSUP's) Fleet and Industrial Supply Center (FISC), Norfolk Detachment Washington. Harman served as a Supply Corps Officer in the Navy from 1985–1992 and retired from the Naval Reserves. She served onboard the USS *Emory S. Land* (AS-39) and earned the Supply Corps Surface Warfare pin. Her other duty stations include: Supreme Allied Command Atlantic, Commander in Chief U.S. Atlantic Fleet, United States Naval Academy, and FISC Norfolk Detachment Washington.

Harman is a member of the DoD Acquisition Professional Community and is Level III certified in Contracting. A Certified Professional Contracts Manager through the National Contract Management Association, she holds a Bachelor of Science degree in Physical Science from the United States Naval Academy, and a master's degree in Management/Acquisition and Contract Management from the Florida Institute of Technology. Harman is a member of Leadership Southern Maryland's Class of 2010.

Harman is a graduate of NAVSUP's Corporate Management Development Program, NAVAIR's Senior Executive Leadership Development Program, and the Federal Executive Institute. Harman has a number or personal and command decorations including the DoN's Meritorious Civilian Service Medal, DoN's FY2010 Acquisition Excellence Award, and the 2015 Public Servant Award from the St. Mary's County Chamber of Commerce.

Examining the Effects of Set Aside Policies on Competition and Growth for Small and Mid-Sized Suppliers[1]

Trevor L. Brown—is a Professor and Dean of the John Glenn College of Public Affairs at The Ohio State University. Brown's research on public procurement and contract management has been published in a variety of peer-reviewed academic outlets (e.g., Cambridge University Press, *Journal of Policy Analysis and Manag*ement). He has also produced numerous consulting and project reports for the U.S. Navy, the Pew Center on the States, and the IBM Center for the Business of Government. Brown served for 20 years in a variety of contract management capacities for a technical assistance contract with the U.S. Agency for International Development. [brown.2296@osu.edu]

Amanda M. Girth—is an Assistant Professor at the John Glenn College of Public Affairs at The Ohio State University. Her research focuses on government contracting with a specific interest in performance and accountability, which has been published in the *Journal of Public Administration Research and Theory*, *Journal of Supply Chain Management, Public Administration Review,* and other outlets. Previously, Girth was a manager for a global consulting firm where she supported information technology initiatives at the U.S. Department of State and U.S. Agency for International Development. She also served in Michigan state government. [girth.1@osu.edu]

Abstract

We examine the federal small business set aside program and assess the impact of small business set asides on supplier competitiveness, program participation, and firm growth. We track the performance of over 700 small businesses over a 10-year period (FY2005–FY2014). We analyze firm-level characteristics and attributes of their federal contracting portfolios. Our exploratory study tests hypotheses to (a) determine whether there is a difference between firms that remain a small business throughout the 10-year period and those that transition to the middle market and become mid-sized firms, and (b) establish a framework for future study.

Our preliminary results show that there are differences between firms that remain a small business during the 10-year period and those that grow into the middle market. Firm attributes that differ include whether the business is woman-owned, the creditworthiness of the firm, firm efficiency, and the firm's number of corporate relationships. Federal portfolios are also different between the two groups. Firms that grow into the middle market on average have contracts in more agencies, across more product or service lines, and have more contract actions related to multi-award vehicles.

Introduction

Small firms benefit from set asides and other programs offered by the U.S. Small Business Administration (SBA). Alternatively, large companies have internal capacity, scale, and extensive past performance history to compete in the public procurement market by bringing financial, personnel, and political resources to bear to win and execute contracts. Mid-sized firms are essentially left out—they are too big to qualify for set asides, yet do not have parity with large firms against whom they are competing for procurements. Anecdotal

[1] Draft. Do not cite.

evidence of this disparity exists (perhaps best underscored by the work of trade associations such as the Association for Corporate Growth, Mid-tier Advocacy, GTSC-Lion's Den, and the development of the bi-partisan Congressional Caucus for Middle Market Growth). However, there is a dearth of empirical evidence on both the structural barriers that exist for middle market firms and the effects of their competitive disadvantage. In the absence of empirical study, notions and rhetoric prevail. This study begins to clarify these claims by analyzing the contours of the competitive federal procurement market for small and mid-sized suppliers.

Inequities in the public procurement market are not an insignificant concern. The scale and scope of the federal procurement market is vast, with over 5,000 different types of products procured (Brown, 2013), and over $438 billion in contracts obligated in 2015 (accounting for approximately 2.5% of GDP). "Middle market" firms account for one-third of private sector GDP and one-third of U.S. jobs. (We use the National Center for the Middle Market definition of "middle market," firms with annual revenues between $10 million and $1 billion.) However, it remains unclear whether mid-sized firms are correspondingly represented in the federal procurement market. A recent study by the Center for Strategic & International Studies (CSIS) suggests the answer is no (Ellman, Morrow, and Sanders, 2011). The CSIS study found that mid-size market share of federal professional services contracts is shrinking. Mid-sized contractors claimed 40% of the total value of federal professional services contracts in 1995, but only 30% in 2009. During the same time period, large contractors have increased their market share from 41% to 48%, and small business market share increased from 19% to 22%. Understanding the barriers to competition, purported disparities, and structural policy effects that impede the middle market firms' ability to compete for federal contracts will have a significant impact on understanding their capacity to capture market share, grow business, and deliver value to federal agencies.

Our study of the supply-side of public procurement builds on extant literature undertaken to develop and test theories of the demand-side of public procurement. We examine the federal small business set aside program and assess the impact of small business set asides on supplier competitiveness, program participation, and firm growth. As there is little direct empirical evidence to draw upon in crafting the study, we test hypotheses to establish an agenda for future research in this exploratory study. Specifically, we ask whether there is a difference between firms that remain a small business throughout the 10-year period and those that transition to the middle market. To do this, we track the performance of over 700 small businesses over a 10-year period (FY2005–FY2014). We analyze firm-level characteristics and attributes of their federal contracting portfolios. The sample of firms is randomly selected and includes firms with contracts for products varying in complexity, from simple product procurements to more complex services contracts (e.g., IT systems).

We postulate that there are unique drivers for firms that grow into the middle market because the incentives are so strong to remain a small business, even if that means stymied growth. We expect that firms that successfully transition out of the small business market have unique ways of overcoming the "benefit cliff" they encounter as they grow. We consider whether current policies governing procurement hamper mid-sized firm competitiveness in the federal procurement market and dampen U.S. economic growth.

Our preliminary statistical tests reveal differences between firms that remain a small business during the 10-year period and those that grow into the middle market. Firm attributes that differ include whether the business is woman-owned, creditworthiness, firm efficiency, and the number of corporate relationships. Federal portfolios are also different between the two groups. Firms that grow into the middle market on average have contracts in more agencies and across more product or service lines, and have more contract actions related to multi-award vehicles.

Our paper proceeds as follows. First, we provide context for the study by describing the federal policy environment for small and mid-sized suppliers. We synthesize two lines of inquiry—contracting and external market dynamics—to build a framework for assessing the impact of inclusion policies and market conditions on small and mid-sized businesses and federal purchasing agencies. Second, we outline findings from the management and entrepreneurship literature on the growth drivers for small and mid-sized enterprises. Next, we introduce the research design of our broader research project, then present the data, measures, and method of this study. We then offer preliminary findings from our statistical analysis. We conclude by discussing potential alternatives to explain the findings and address wider implications of the study.

Set Aside Policies in Federal Procurement

We begin by considering initiating legislative action that established small business guidelines and the Small Business Administration (SBA). The first substantive guidance directed to federal agencies to contract with small businesses originated in the U.S. Senate in 1940 with the Special Committee to Study and Survey Problems of Small Business Enterprises, and in the U.S. House of Representatives in 1941 with the Select Committee on Small Business. The Committees were created to protect the interests of small business owners, recognizing the need for a thriving small business community for innovation, economic growth, and national security. The Small Business Act of 1953 explicitly stated that government prime contracts and subcontracts should be awarded to small business, and later the Small Business Act of 1958 created the SBA, an independent agency within the Executive branch. In 1975, the Congressional Committee on Small Business was made a permanent standing committee to, in part, oversee the SBA. The Senate followed suit in 1981 when it created the standing Committee on Small Business, which in 2001 was renamed the Committee on Small Business and Entrepreneurship (Office of Small Business Programs, n.d.; Senate Committee on Small Business & Entrepreneurship, n.d.; Small Business Committee, n.d.).

The SBA establishes overall and agency-specific procurement goals. One federal procurement goal is statutorily established as 23% of contract value of prime contracts for small businesses. There are goals within that subset, such as 5% of prime and subcontracts awarded to woman-owned small businesses and 5% of prime and subcontracts to small disadvantaged businesses, among others. Agencies biennially negotiate their targets with the Small Business Administration in order to meet government-wide goals. In FY17, goals ranged from 10% at the Department of Energy to 73% at SBA. Additionally, federal agencies set annual goals for subcontracts. For example, the Department of Defense’s prime contract goal is 22% in FY17, but the subcontracting goal is 34%.

Small business procurement policies can be viewed as largely “policy ambivalent,” perhaps best illustrated by preferences toward supporting small and mid-sized firms through aspirational, goal-oriented policies rather than enforcement (Kidalov & Snider, 2011). Current approaches to meet procurement goals rely heavily on administrative discretion, and yet as Snider, Kidalov, and Rendon (2013) found, discretion has been considerably reduced in recent years.

In practice, acquisition officials are asked to deliver contracts that meet best value, low cost, or other performance objectives, and to meet broader political objectives that can affect (constrain) eligible suppliers. Public sector contracts are not simply a tool to increase efficiency; they can also serve to promote other public values. Procurement policies that target specialized groups, such as small businesses, minority-owned, or women-owned firms, are designed to promote equity and representativeness. As with most government policies, unintended adverse effects can result in pursuit of overcoming market failures (Vining & Weimer, 2005). Whereas competition is a basic assumption underlying public sector procurement, procurement policy and regulation favoring small businesses restrict competition and contribute to weakly competitive procurements (Girth et al., 2012). When markets are constrained, purchasers have fewer choices to balance different, and sometimes competing, purchasing goals (Brown, Potoski, & Van Slyke, 2013; Johnston & Girth, 2012). Set aside policies can further constrain market competition to a narrow pool of suppliers and thereby limit the range of cost, quality, and delivery options for goods and services procured under said programs (Brown, 2007). Despite the prevalence and importance of inclusion policies in federal procurement, very little is known about the design or impact of these programs on purchasing agencies or the supplier market.

Growth Drivers: Small and Mid-Sized Businesses

Set asides and supplier diversity efforts are not limited to public procurements; rather, there are competitive advantages in the private market for diverse supplier representation (Richard, 2000). Small businesses have advantages over large firms in that they are more responsive and innovative, in part because they can be more swift and flexible (Dean et al., 1998). Yet as Dobbs and Hamilton's (2007) extensive review of the literature demonstrated, there is no single deterministic factor in predicting firm growth. They also concluded that scholarly understanding of growth for small and mid-sized enterprises is inadequate.

Research on the growth of small and medium-sized enterprises has focused on managerial strategies, leader characteristics, environmental factors, and firm attributes such as human resource, organizational, marketing, and financial capabilities (Barbero, Casillas, & Feldman, 2011; Dobbs & Hamilton, 2007). There are factors within each of these categories that have significant explanatory effects on small business growth. First, a number of studies cite the positive impact of the availability of resources and external financing on firm growth (Becchetti & Trovato, 2002). Lack of cash flow and access to external financing can hinder growth by limiting availability to manage operations and capturing strategic market opportunities (Carter & Van Auken, 2005; Locke, 2005). Younger firms also tend to experience more rapid growth (Lotti, Santarelli, & Vivarelli, 2003). While firm size, whether number of employees or revenues, is a measure of growth, it can also be a determinant of growth rates (Evans, 1987).

Owner motivation (and then lack of motivation to grow once the owner has reached sufficient income) can drive growth trajectories for small businesses (Robson & Bennett, 2000). Leadership characteristics are also a determinant of firm growth, culturally establishing a growth orientation (Barringer & Jones, 2004). Although as Smallbone, Leig, and North (1995) found, developing internal capacity to allow for strategic leadership is an important factor in small business growth. In addition to internal strategy, externally-facing actions such as developing collaborative relationships through trade associations, lobbying and other external alliances can precipitate growth (Robson & Bennett, 2000).

Finally, pursuing differentiation strategies appear to correlate to small business growth. Small businesses do this by actively managing their products and markets (Smallbone et al., 1995) and innovating to compete in markets with larger firms (O'Gorman, 2001). In sum, there are a number of factors that emerge in the management and entrepreneurship literature that predict small business growth. We assess a number of these factors in our study by testing differences among small businesses operating in the federal small business set aside market.

Data and Method

Research Design

We examine the performance of small businesses contracting with the federal government by randomly selecting 1,025 businesses that have taken advantage of the small business set aside program in 2005 (i.e., they have at least one contract action associated with a small business set aside prime contract). We stratify the sample such that 60% of contracts are Department of Defense (DoD) contracts to mirror federal spending. The sample includes firms with contracts for products and services varying in complexity, from simple product procurements to more complex services contracts such as IT systems.

The unit of analysis is firm-year. This permits us to analyze the annual performance of sampled firms over a 10-year period, and includes firm attributes and their federal contracting portfolio.

Our methodological approach is descriptive, as we begin to understand the differences between small businesses that strategically stay small and those that attempt to grow beyond the small business market. This research is our first step in a large research project aimed at disentangling the procurement environment for small and mid-sized suppliers.

Data

Data is gathered from two sources: Federal Procurement Data System-Next Generation (FPDS-NG) and Dun & Bradstreet. Contracts data for each of the sampled firms is compiled from FPDS-NG. Data on firm attributes is procured from Dun & Bradstreet. The unit of analysis is firm-year. Contracts data from FPDS-NG is aggregated to account for contract activity for each fiscal year. Dun & Bradstreet data is reported annually. The data sets are then merged by firm-year. The process for cleaning, coding, and merging the data sets is documented in the appendix.

Some of the firms that have contract actions categorized as a small business set asides in 2005 already “outgrew” their small business status. That is, they no longer self-certified as a small business by the SBA. (Firms self-certify through the SBA and their revenue and/or employee thresholds are reported on a three-year rolling average.) These firms are excluded from our present study.

The purpose of this analysis is to describe the difference between firms that were small businesses throughout the 10-year period, and those that outgrew their small business status at some point over that time period. Those that reverted back to small were excluded from this study. As such, this dataset includes 721 firms that started as small businesses in 2005 and remained small businesses through 2014, and 46 firms that started as a small business in 2005 but grew into the middle market at some point during the 10-year period. As observations are firm-year, we have a total of 7,670 observations in the dataset.

Measures

Our dependent variable, *small business*, is binary with 0 representing small businesses that started small in 2005 and stayed small through 2014 (n=7,210) and 1 representing firms that started small in 2005 and grew into the middle market at some point during the time period (n=460). Dun & Bradstreet retrieves data for this variable directly from the SBA.

Our independent variables include firm characteristics and federal contracting portfolio attributes, which are listed below. Descriptive statistics are found in Table 1.

Firm Characteristics

- *Credit rating*. Credit worthiness scored by Dun & Bradstreet analyst. 1=limited, 2=fair, 3=good, 4=high. Source: Dun & Bradstreet.
- *Minority ownership*. Indicates whether a minority owns a majority of the business. 1=minority-owned, 0=not minority-owned. Source: Dun & Bradstreet.
- *Woman ownership*. Indicates whether a woman owns a majority of the business. 1=woman-owned, 0=not woman-owned. Source: Dun & Bradstreet.
- *Firm age*. Contract year (2005–2014) less the year business started. Source: Dun & Bradstreet.
- *Lines of business*. Number of lines of businesses in which the organization is engaged. Source: Dun & Bradstreet.
- *Number of DUNS family members*. Number of corporate family relationships identified by Dun & Bradstreet. Source: Dun & Bradstreet.
- *Number of employees*. Total number of employees in the organization. Source: Dun & Bradstreet.
- *Annual Sales.* Total annual sales volume. Source: Dun & Bradstreet.
- *Efficiency*. Value to determine efficient use of labor resources. Derived by dividing total annual sales volume by the number of employees.

Federal Contracting Portfolio

- *Agency diversity*. Count of the number of federal agencies the firm has contracts with for a contract year. Source: FPDS-NG
- *NAICS diversity*. Count of the number of different NAICS the firm's contracts are specified for a contract year. NAICS are aggregated to the first two digits (e.g., naics53, naics54). Source: FPDS-NG
- *PSC diversity*. Count of the number of different PSCs the firm's contracts are specified for a contract year. PSCs are aggregated to the first two digits for products (e.g, psc70, psc71) and the first letter for services (e.g., psca, pscb). Source: FPDS-NG
- *Contract actions associated with IDV*. Number of contract actions that are associated with an indefinite delivery vehicle (IDV) for a contract year. Source: FPDS-NG

Table 1. Descriptive Statistics

Variable	Obs	Mean	St dev	Min	Max
Small business	7,670	0.599	0.237	0	1
Credit rating	4,942	2.463	0.687	1	4
Minority ownership	7,249	0.165	0.372	0	1
Woman ownership	7,249	0.264	0.442	0	1
Firm age	6,902	25.12	20.01	0	214
Lines of business	7,249	1.452	0.849	1	6
No. DUNS family (log)	7,249	0.228	0.797	0	8.62
No. of employees (log)	7,249	2.423	1.232	0	6.55
Annual sales (log)	7,249	13.46	2.442	0	19.2
Efficiency	7,249	7.032	4.087	0	21.5
Agency diversity	4,384	1.978	2.712	1	40
NAICS diversity	4,384	1.574	1.039	1	10
PSC diversity	4,384	2.450	2.727	1	24
Contract actions – IDV (log)	4,384	0.846	1.292	0	7.56

[1] Observations vary due to missing values.

Method of Analysis

As we are undertaking an exploratory study, we descriptively analyze two groups by performing difference of means tests across a number of theoretically relevant variables. Across all cases, *small business* is the dependent variable. We ask: Is there a difference between firms that remain a small business throughout the 10-year period and those that transition to the middle market? We use appropriate parametric and non-parametric testing (t-test and chi-square) based on the distribution of the independent variable.

Results

Hypothesis testing reveals that there are differences between the firms in our sample that remained small businesses and those that grew into the middle market. We find key differences in firm characteristics and in their federal contracting portfolio. The results of our analysis are reported in Table 2.

Among the firm characteristics we measure, we find that firms that outgrew the program and emerged into the middle market were less likely to be woman-owned firms; however there was no significant difference among firms that were minority-owned. For both women-owned and minority-owned firms, there are additional set aside goals that benefit these firms, potentially making the transition to the middle market even less attractive.

The firms that emerged into the middle market have lower average credit ratings. This is a curious finding, considering that small business growth is found to be tied to access to external findings, and creditworthiness is a key predictor. These firms are also markedly less efficient (4.86 logged) than the firms that remained small (7.48 logged), when we measure efficiency as a function of sales revenue and number of employees. There also might also be greater risk tolerance among the firms that emerge in to the middle market, as indicated by the slightly lower credit scores, and lower efficiency measures. Another explanation for lower efficiency measures is the investment in internal infrastructure to facilitate growth, which would in turn depress the efficiency measure we utilize. We find differences in annual sales and number of employees between the two samples, which is expected and serves as a validity check, as those are the factors that drive qualification for the SBA program.

The other striking difference between the two groups in the sample is the number of DUNS family members. Firms that grow into the middle market have a far higher average number of DUNS family members (1.78 logged) than those that remain a small business (0.12 logged). This likely indicates that some of the firms that emerge into the middle market have been acquired.

Analyzing features of the firms' federal contracting portfolios, we find that firms that grow into the middle market have greater agency diversity. That is, they have greater breadth across the federal government, operating on average in 2.77 different agencies as compared to firms remaining a small business which operated on average in 1.99 agencies. We also observe that firms that emerge into the middle market were more likely to have contracts classified in more NAICS codes (average of 1.89 compared to 1.55) and PSCs (average of 23.4 compared to 2.38) than those that remained a small business. This might mean that these firms engage in different functional areas, or that they are adept at navigating thresholds for varying NAICS codes in accordance with their SBA self-certification.

Firms that emerge into the middle market are also more likely to have a greater number of contract actions associated with multiple-award contracts. Firms emerging into the middle market have an average of 1.07 IDV-related contract actions compared to an average of 0.83 among small businesses (values are logged).

Table 2. Results of Statistical Tests

Variable	Small business (0) Obs	Mean	St er	Outgrew (1) Obs	Mean[1]	St er	t-stat
Credit rating[2]	4,593	2.47	0.10	349	2.35*	0.03	2.968
Minority ownership[2]	6,792	0.16	0.00	457	0.15	0.01	0.353
Woman ownership[2]	6,792	0.27	0.01	457	0.12*	0.02	7.235
Firm age	6.479	25.1	0.25	423	24.5	1.05	0.683
Lines of business	6,792	1.45	0.10	457	1.47	0.38	-0.513
No. DUNS family (log)	6,792	0.12	0.00	457	1.78*	0.11	-49.96
No. of employees (log)	6,792	2.34	0.14	457	3.55*	0.06	-2.017
Annual sales (log)	6,792	13.5	0.03	457	13.7*	0.22	-49.96
Efficiency (log)	6,792	7.48	0.49	445	4.86*	0.16	13.23
Agency diversity	4,070	1.91	0.04	314	2.77*	0.20	-5.391
NAICS diversity	4,070	1.55	0.02	314	1.89*	0.07	-5.741
PSC diversity	4,070	2.38	0.04	314	23.4*	0.21	-6.264
Contract actions – IDV (log)	4,070	0.83	0.02	314	1.07*	0.08	-3.161

[1] *p<0.00, Two-tailed tests

[2] *t*-test results reported are ease of reporting, chi-square analysis performed and results consistent with *t*-test

We have not included measures on competition in this analysis. We suspect that there will be differences among the two groups related to their competitive positioning in the federal market and their use of the small business set asides. We intend to include these measures in future studies.

Discussion

Our preliminary results yield interesting findings on the differences between the two groups of small businesses in this study. Small businesses that grow into the middle market are not just unique in their sales volume and number of employees, but also in financial and managerial aspects. Their federal contracting portfolios yield higher levels of activity in federal markets, across product lines, and in multiple-award contracts. These findings indicate further exploration is needed to discern the impact of each of these factors, among others, on emergence into the middle market. In this section, we explore some of the possible explanations for the patterns we observe. We also discuss future research opportunities in light of our exploratory study.

Our interest in this research lies in the intent of the small business set aside program and the practical implications of the constrained competitive environment. The vast majority of firms in our sample shelter in the small business set aside market. They fail to grow beyond the sales or employee thresholds in the product or service areas they have self-certified in. In most cases, firms in our sample elect to stay small. These firms recognize the value of the constrained federal market established for small businesses. Their clients also value their small business status, allowing for more desirable procurements as they help to achieve the agency's small business goals. In other cases, we suspect small businesses are unable to harness the resources, whether financial or managerial, to grow.

When firms make strategic decisions to stay small in order to retain small business status for federal procurements, they artificially constrain their growth in order to stay under the revenue or employee thresholds for specified NAICS. If the intent of federal policies to support small business is to encourage economic growth and innovation, then firm behavior does not necessarily align with these goals. The behavior, albeit rational on the small businesses' part, is an unintended consequence of creating markets and subsidizing subsets of industry.

To that end, there are 77 firms in our larger dataset that moved in and out of the small business market two or more times during the 10-year time period we studied (but that were not included in this study). These firms started as small businesses, grew into the middle market, and then re-certified as a small business after at least one year. We suspect there are also different attributes that drive this behavior among firms in this group and intend to explore this in future research. Yet this phenomenon reinforces questions about the behaviors of firms that are not able to thrive as they emerge into the middle market, and either intentionally constrain to fall meet small business thresholds in later years or fail to win contracts when competing outside of the set aside market.

In addition to the future work already identified, we plan to use interviews to qualitatively explore the varying strategic positioning by firms. In particular, to what extent do perverse incentives to curb growth impact broader economic growth? Interviews will also help us to identify other strategies not gleaned from these data that impede or support firm growth into the middle market.

Should some of these relationships hold as we apply more advanced econometric models to these data, several policy implications might bear out. These include strategies to support lasting transition to the middle market. While there is little drive to create additional set aside categories aimed to benefit mid-sized firms (particularly those at the lower threshold of the middle market), it remains a policy option to carve out a niche in the small business space to support mid-sized transition. There are also other ways the federal government could support suppliers in the middle that are neither large nor small. Creating federal supply schedules for mid-sized suppliers is one alternative that has been advanced by trade associations supporting mid-tier suppliers.

Another key issue for policymakers to consider is the way in which procurement policy and regulation treat federal suppliers. Procurement policy and Federal Acquisitions Regulation (FAR) recognize only two categories of suppliers: small business and not small. Small firms experience a significant "benefit cliff" as they grow out of the small business market. Treating mid-sized firms, particularly those at the lower threshold of the market, as large firms might be adversely affecting public value. On the one hand, a mid-sized firm may be leaner and be able to aggressively price proposals to compete with larger firms. On the other hand, mid-sized firms do not have robust internal resources or past-performance qualifications to technically compete for procurements, compared to large firms. Under this scenario, the government may be losing an opportunity to secure value by not actively seeking contracts with mid-sized firms. It is understood that small businesses are higher risk suppliers. Unlike large firms, their internal processes are immature, and resource shocks can have profound effects on a small enterprise and their clients. Large firms largely mitigate that risk, but can be costly. Mid-sized firms can be well suited to provide value at lower risk than small firms and lower cost than large. In either case, federal agencies are likely missing opportunities to secure value by treating all firms that are not small as large and not incentivizing contracting with mid-sized suppliers.

Conclusion

This study represents the first step in a broader research project on the effects of set aside policies on small and mid-sized federal suppliers. The findings reveal discernible patterns among firms that grow into the middle market after benefiting from small business set aside programs. Our preliminary results show that there are firm characteristics that might drive emergence into the middle market, to include whether the firm is woman-owned. Small businesses that grow into the middle market on average have lower credit scores and lower efficiency measures. There are also more formal corporate relationships—likely an indicator of acquisition—among firms that emerged in the middle market. Federal portfolios are also different between the two groups. Firms that grow into the middle market on average have contracts in more agencies, across more NAICS and PSCs. They also have a higher number of contract actions associated with IDVs.

The results raise questions for further study. Our next steps are to determine whether these patterns hold as we perform more advanced econometric analysis to predict membership in the two groups. Future research to include qualitative analysis is also warranted. As such, we are in the process of interviewing suppliers to determine their strategies for success, and federal government officials to obtain their perspectives on the small business set aside program and mid-sized suppliers.

Our study is a critical first step in capturing the structural dynamics involved in the design, implementation, and evaluation of competitive practices in federal agencies aimed at promoting small business participation and growth. For small and mid-sized businesses, our preliminary results provide empirical evidence of the differences between firms that have grown into the middle market and those that have not. The results are promising for informing the drivers of growth for small and mid-sized firms, and the strategies that enable successful firm development. The results also have implications for policy makers—if further analysis shows that middle market firms are, on balance, unable to compete in the federal procurement market, then agencies are likely missing critical opportunities to secure value.

References

Barbero, J. L., Casillas, J. C., & Feldman, H.D. (2011). Managerial capabilities and paths to growth as determinants of high-growth small and medium-sized enterprises. *International Small Business Journal, 29*(6), 671–694.

Barringer, B. R., & Jones, F. F. (2004). Achieving rapid growth: Revisiting the managerial capacity problem. *Journal of Developmental Entrepreneurship, 9*(1), 73.

Becchetti, L., & Trovato, G. (2002). The determinants of growth for small and medium sized firms: The role of the availability of external finance. Small Business *Economics, 19*(4), 291–306.

Brown, T. L. (2007). Testimony on the Small Business Administration's Contracting Programs, Small Business Committee, U.S. House of Representatives, Washington, DC, September 19, 2007.

Brown, T. L. (2013). *A guide for agency leaders on federal acquisition: Major challenges facing government*. Washington, DC: IBM Center for the Business of Government.

Brown, T. L., Potoski, M. & Van Slyke, D. M. (2013). *Complex contracting: Government purchasing in the wake of the U.S. Coast Guard's Deepwater Program*. New York, NY: Cambridge University Press.

Carter, R. B., & Van Auken, H. (2005). Bootstrap financing and owners' perceptions of their business constraints and opportunities. *Entrepreneurship & Regional Development, 17*(2), 129–144.

Dean, T. J., Brown, R. L. & Bamford, C. E. (1998). Differences in large and small firm responses to environmental context. *Strategic Management Journal, 19*, 709–728.

Dobbs, M., & Hamilton, R. T. (2007). Small business growth: Recent evidence and new directions. *International Journal of Entrepreneurial Behavior & Research, 13*(5), 296–322.

Ellman, J., Morrow, D., & Sanders, D. (2011). *Structure and dynamics of the U.S. federal professional services industrial base.* Washington, DC: Center for Strategic & International Studies.

Evans, D. S. (1987). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *Journal of Industrial Economics*, 567–581.

Girth, A. M., Hefetz A., Johnston, J. M., & Warner, M.E. (2012). Outsourcing public service delivery: Management responses in noncompetitive markets. *Public Administration Review, 72*(6), 887–900.

Johnston, J. M., & Girth, A. M. (2012). Government contracts and "managing the market": Exploring the costs of strategic management responses to weak vendor competition. *Administration & Society, 44*(1), 3–29.

Kidalov, M. V,. & Snider, K. F. (2011). U.S. and European public procurement policies for small and medium-sized enterprises (SME): A comparative perspective. *Business and Politics, 13*(4), 1–41.

Lotti, F., Santarelli, E., & Vivarelli, M. (2003). Does Gibrat's Law hold among young, small firms? *Journal of Evolutionary Economics, 13*(3), 213–235.

Office of Small Business Programs. (n.d.). Legislation: A brief history of key small business acts. Retrieved April 1, 2017, from http://business.defense.gov/legislation/

O'Gorman, C. (2001). The sustainability of growth in small-and medium-sized enterprises. *International Journal of Entrepreneurial Behavior & Research, 7*(2), 60–75.

Richard, O. C. (2000). Racial diversity, business strategy, and firm performance: A resource-based view. *Academy of Management Journal, 43*(2), 164–177.

Robson, P. J., & Bennett, R. J. (2000). SME growth: The relationship with business advice and external collaboration. *Small Business Economics, 15*(3), 193–208.

Senate Committee on Small Business & Entrepreneurship. (n.d.). History. Retrieved April 1, 2017, from https://www.sbc.senate.gov/public/index.cfm?p=History

Smallbone, D., Leig, R., & North, D. (1995). The characteristics and strategies of high growth SMEs. *International Journal of Entrepreneurial Behavior & Research, 1*(3), 44–62.

Snider, K. F., Kidalov, M. V., & Rendon, R. G. (2013). Diversity governance by convenience? Federal contracting for minority-owned small businesses. *Public Administration Quarterly, 37*(3), 393–432.

U.S. Small Business Committee. (n.d.). Committee history. Retrieved April 1, 2017, from http://smallbusiness.house.gov/about/

Vining, A., & Weimer, D. (2005). Economic perspectives on public organizations. In D. Ferrin, L. Lynn, & M. Pollitt (Eds.), *Oxford handbook of public management* (pp. 209–233). Oxford, England: Oxford University Press.

Acknowledgments

This material is based upon work supported by the Naval Postgraduate School Acquisition Research Program under Grant No. N00244-16-1-0049. The views expressed in written materials or publications, and/or made by speakers, moderators, and presenters, do not necessarily reflect the official policies of the Naval Postgraduate School nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. government.

Professors Girth and Brown are Fellows of the National Center for the Middle Market at the Fisher College of Business, OSU, and acknowledge the Center's support for this research.

Appendix

Selection

Criteria for selection was whether a firm had a contract action in 2005 that was designated a small business contract. Of the 1,025 firms selected, 143 firms had outgrown small business status in the years preceding 2005 but still had active contracts that were initially awarded as small business contracts.

FPDS Aggregation

FPDS-NG reports all contract actions for unclassified contracts over $3,000 in value. We aggregated all contract actions for a given year for each firm so that we have the firm-year as the observation.

Missing Data

In some cases, contract actions reported in FPDS-NG were missing key data elements. We dropped contract actions from the analysis if they were missing the following information: contract pricing type, product or service code, principal NAICS code, or contracting agency. These contract actions were removed from the data set prior to aggregating the contract actions.

There were also instances where Dun & Bradstreet data was missing for a particular year. In cases where we were confident the firm continued to exist, we imputed the missing data using linear interpolation for continuous variables and modal imputation for nominal variables. Our imputation procedures were calculated for the unique firm. That is, if one year of Dun & Bradstreet data was missing, the other nine years of firm data was used to calculate the imputed values. There were other cases where we elected not to impute missing data because we were not confident the firm continued to exist. For those firm/year entries, we were either also missing corresponding contracts data for that year and/or we were missing contracts and Dun & Bradstreet data for years following, thus making it impossible to determine whether the firm was in business or not.

In no instance did we impute FPDS-NG data. In some cases, we have missing contracts data that do not allow for complete analysis of firm performance in the federal contracting market in a given year, which accounts for the variation in the number of observations across variables of interest.

Evaluating the Impact of Small Business Set-Asides on Acquisitions Efficiency

William Lucyshyn—is the Director of Research and a Research Professor at the Center for Public Policy and Private Enterprise in the School of Public Policy, University of Maryland. In this position, he directs research on critical policy issues related to the increasingly complex problems associated with improving public-sector management and operations, and with how government works with private enterprise. Current projects include modernizing government supply-chain management, identifying government sourcing and acquisition best practices, and analyzing Department of Defense business modernization and transformation. Previously, Lucyshyn served as a Program Manager and the Principal Technical Advisor to the Director of the Defense Advanced Research Projects Agency (DARPA) on the identification, selection, research, development, and prototype production of advanced technology projects.

Prior to joining DARPA, Lucyshyn completed a 25-year career in the U. S. Air Force. Lucyshyn received his bachelor's degree in Engineering Science from the City University of New York and earned his master's degree in Nuclear Engineering from the Air Force Institute of Technology. He has authored numerous reports, book chapters, and journal articles. [lucyshyn@umd.edu]

John Rigilano—is a Faculty Research Associate at the Center for Public Policy and Private Enterprise. He received his Master of Public Policy degree from the University of Maryland's School of Public Policy in 2011, where he also served as a graduate research associate at CPPPE. Prior to his time at UMD, Rigilano received his Bachelor of Arts from Penn State University in 2004 with a major in Anthropology. During his time at CPPPE, he has coauthored and contributed to numerous reports.[jprig@umd.edu]

Abstract

The Federal Government spent over $470 billion on procurement in FY 2016. Spending of this magnitude creates opportunities for implementing selected national policies. For instance, current law requires that low-cost acquisitions be reserved exclusively for small business concerns, with qualifying businesses assuming the role of prime contractor. However, the pursuit of admirable social goals such as this may not be rational from an economic or technical standpoint.

This report analyzes the distribution of small business procurement across industry sectors using data from the Federal Procurement Data System (FPDS). We show that a relatively small number of large firms dominate the federal contracting landscape in certain sectors, such as defense, and account for a significant proportion of procurement spending. Accordingly, set-aside policy has a disparate impact on the remainder of the spending, concentrating it into certain industry sectors where there are greater opportunities for small businesses, limiting free and open competition, and creating a series of unintended consequences for government (e.g., contracting and economic inefficiency) and small businesses (e.g., uneven and unsustainable growth and barriers to entry into the federal contracting space).

Introduction

The Federal Government, on average, spent half a trillion dollars annually on procurement over the last decade ($470 billion in FY 2016), roughly 40% more than what was spent in real terms during the 1990s (see Figure 1).

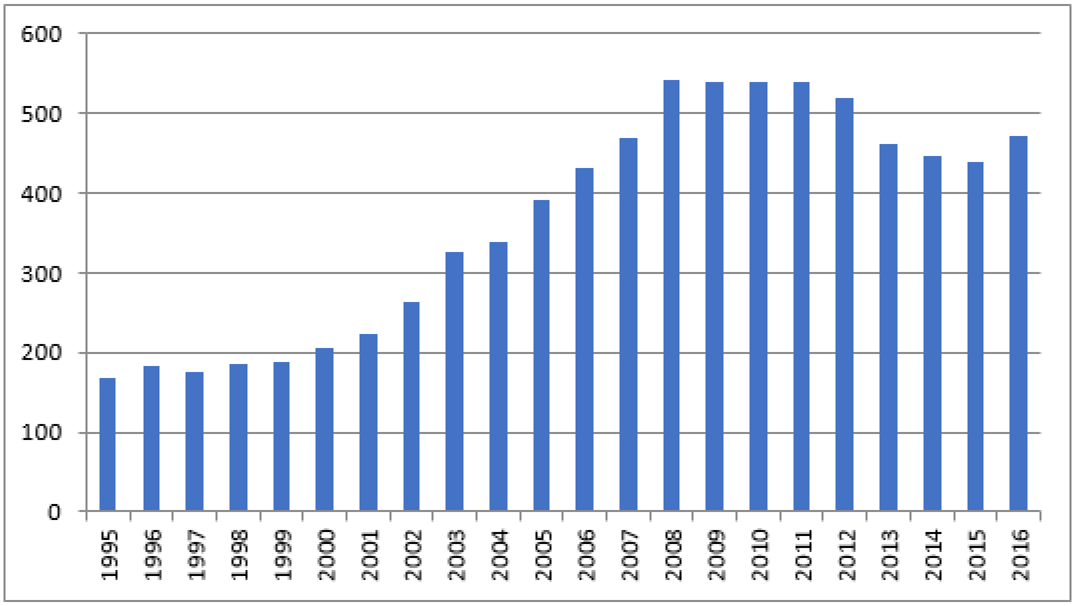


Figure 1. Federal Contract Spending, Action Obligations in $Billions, 1995–2016
(Analysis of FPDS Data)

Spending of this magnitude creates opportunities for implementing socio-economic policies aimed at promoting small businesses, especially those owned by members of historically-disadvantaged groups (i.e., minorities and women). In 1988, Congress began requiring that "the President shall annually establish Government-wide goals for procurement contracts" at specified minimum percentages of procurement (Beale, 2014). The initial government-wide goal for small business procurement was set "at not less than 20 percent of the total value of all prime contract awards." In 1997, the goal was raised to 23%.

As part of these broad socio-economic goals, the Small Business Act of 1953 established the Small Business Administration (SBA) to "aid, counsel, assist, and protect the interests of small business concerns, to preserve free competitive enterprise, and to maintain and strengthen the overall economy of our nation" (SBA 2014). However, as with any effort to regulate a complex system, there are unintended consequences. It remains unclear whether the current set-aside policy, in its current implementation, represents the best strategies for leveraging the capabilities that small businesses can offer.

This report analyzes the distribution of small businesses procurement across industry sectors using data from the Federal Procurement Data System (FPDS). We show that a relatively small number of large firms dominate the federal contracting landscape in certain sectors, such as defense, and account for a significant proportion of procurement spending. Accordingly, set-aside policy has a disparate impact on the remainder of the spending, concentrating it into certain industry sectors where there are greater opportunities for small businesses, limiting free and open competition, and creating a series of unintended consequences for government (e.g., contracting and economic inefficiency) and small businesses (e.g., uneven and unsustainable growth and barriers to entry into the federal contracting space).

The advantages of small business—innovation and agility—have been recognized for decades. Small business is the "driver and engine of growth" and the "lifeblood of our economy" (Obama, 2014). However, there are indications that current set-aside policies fall short of their intended objectives: promoting the growth and prosperity of small business, improving government acquisitions efficiency, and fostering economic growth.

Background

A small business, to qualify as such under SBA requirements, must meet the following criteria:

- Meets SBA industry-specific size standards;
- Is organized for profit;
- Has a place of business in the United States;
- Operates primarily within the United States or makes a significant contribution to the U.S. economy through payment of taxes or use of American products, materials, or labor;
- Is independently owned and operated; and
- Is not dominant in its field on a national basis. (SBA, 2015)

Current law requires that all acquisitions above the micro-purchase threshold of $3,500 be set aside for small business concerns provided that there is a reasonable expectation that offers from at least two responsible small business concerns will be received at fair market prices. This provision is commonly referred to as the “rule of two.”

Set-Aside Goals

In addition, each year the government sets a government-wide small business prime contracting goal. The current goal is 23%. It also establishes goals for small-disadvantaged businesses, women-owned small businesses, historically-underutilized businesses zones (HUBZone), and service-disabled veteran-owned small businesses. There is also a government-wide small business subcontracting goal and subcontracting goals in each of the aforementioned categories. The 2015 set-aside goals and levels of achievement are shown in Table 1.

Table 1. FY2015 Government-Wide Small Business Procurement Goals and Achievement

(SBA, 2016)

Category	Prime Contracting Goal	Prime Contracting Achievement	Subcontracting Goal	Subcontracting Achievement
Small Business	23.00%	25.75% ($90.7B)	34.03%	31.30%
Women-Owned Small Business	5.00%	5.05% (17.8B)	5.00%	5.05%
Small Disadvantaged Business	5.00%	10.06% ($35.4B)	5.00%	10.06%
Service-Disabled Veteran Owned Small Business	3.00%	3.93% ($13.8B)	3.00%	3.93%
HUBZone	3.00%	1.82% ($6.4B)	3.00%	1.82%

Current law also requires that federal agencies, in collaboration with the SBA, establish their own goals biannually in each of the categories listed in Table 1. The goals vary widely by agency. Prior to finalizing each agency’s goals, the SBA determines whether the goals, in the aggregate, meet or exceed the government-wide statutorily mandated goals in each of the categories. The critic might question why agency goals are subordinated to government-wide goals, rather than used to inform, if not justify, the government-wide targets.

The SBA provides each agency with an annual performance scorecard that lists achievement in each category along with an overall grade, using a methodology that heavily weights prime contracting achievement. An agency's grade is composed of three quantitative measures: prime contracts (80%), subcontracts (10%), and its "progress plan" for meeting future goals (10%; SBA, 2015). Accordingly, comparing their letter grades cannot reveal which agencies relied more heavily on small business to meet their procurement needs. Table 2 shows the DoD's 2016 small business procurement scorecard.

Table 2. DoD's 2016 Small Business Procurement Scorecard

Department of Defense
FY2015 Small Business Procurement Scorecard

A
106.34%

FPDS-NG Prime Contracting Data as of Feb. 20, 2016
eSRS Subcontracting Data as of Mar. 14, 2016

Prime Contracting Achievement:			88.02%
	2014 Achievement	2015 Goal	2015 Achievement
Small Business	23.47%	21.60%	24.64% ($52.4 B)
Women Owned Small Business	3.97%	5.00%	4.43% ($9.4 B)
Small Disadvantaged Business	8.95%	5.00%	9.53% ($20.2 B)
Service Disabled Veteran Owned Small Business	3.04%	3.00%	3.45% ($7.3 B)
HUBZone	1.93%	3.00%	1.87% ($4.0 B)

Subcontracting Achievement:			8.46%
	2014 Achievement	2015 Goal	2015 Achievement
Small Business	33.20%	36.00%	32.30%
Women Owned Small Business	5.70%	5.00%	5.30%
Small Disadvantaged Business	4.60%	5.00%	4.40%
Service Disabled Veteran Owned Small Business	2.10%	3.00%	2.00%
HUBZone	1.50%	3.00%	1.40%

SBA Size Standards

One of the challenges created by the small business set-aside policies is defining what a "small" business is. These definitions are tailored to industry classifications and have evolved over time. The SBA devises size standards which are expressed as either the average number of over the past 12 months or average annual receipts over the past three years. The size standard varies by NAICS industry and is dependent on an SBA methodology that analyzes five primary factors within each industry: average firm size, degree of competition within an industry, startup costs and entry barriers, distribution of firms by size, and small business share in federal contracts.

> ***NAICS Codes***
> The North American Industry Classification System (NAICS) is the standard used by federal statistical agencies, including the SBA, in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the U.S. business economy. NAICS recognizes 20 industry sectors, which are separated into 99 subsectors, 312 industry groups, and over 1000 industries, each of which is assigned a six-digit code.

Contracting officers then must classify each and every solicitation by an industry-level NAICS code that, by their determination, describes the principal purpose of the product or service. Accordingly, a business that qualifies as "small" under one or more NAICS codes may not qualify under others. As one might imagine, the procuring agency must carefully consider each NAICS code designation. Erroneously assigned codes constitute valid ground for bid protests, which can be costly for the government.

However, NAICS code selection can be a subjective endeavor, and can significantly affect the companies eligible. McVay (2009) provides an example which would be comical in its banality if not for its real-world implications. He writes that "if a contracting officer decides to set aside a contract for paperboard boxes, should he categorize the boxes as 'Setup Paperboard Boxes' (NAICS code 322213), which has a size standard of 500 employees, or as 'Folding Paperboard Boxes' (NAICS code 322212), which has a size standard of 750 employees?" (p. 185).

Small Business Representation in Federal Contracting

Figure 2 depicts FY 2015 federal contract obligations by industry sector. The first thing to notice is that federal procurement is highly concentrated by sector, with manufacturing; professional, scientific, and technical services; construction; and administrative support accounting for more than 80% of procurement.

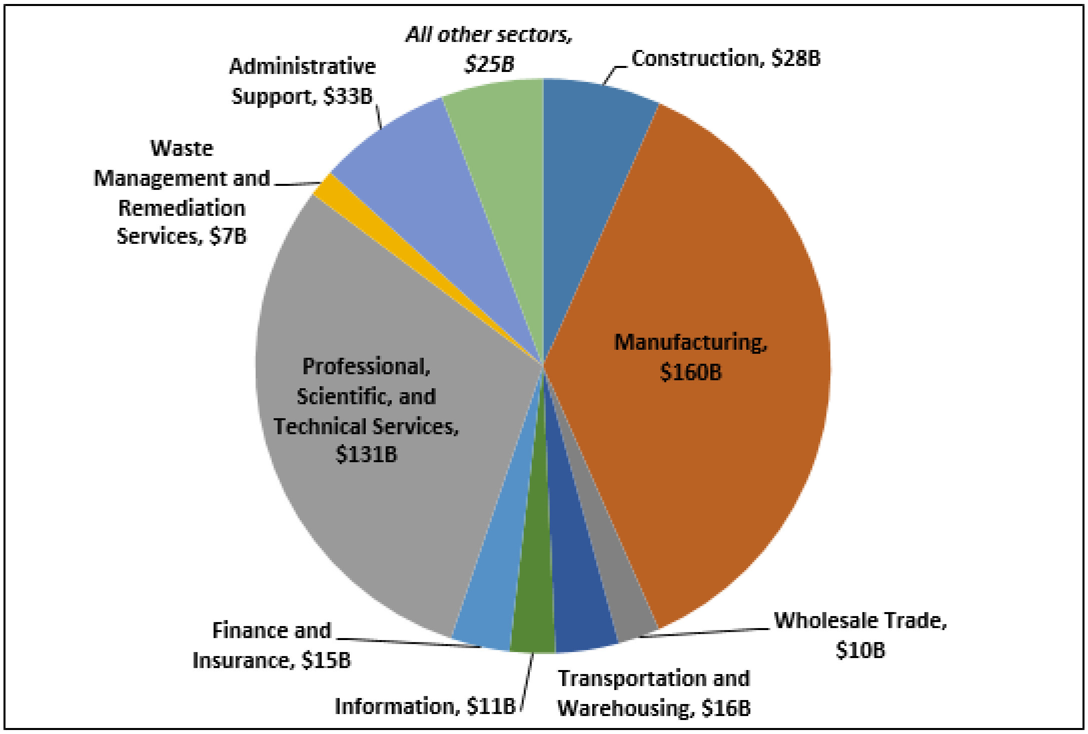


Figure 2. FY 2015 Total Federal Contract Obligations by Industry Sector
(Analysis of FPDS Data)

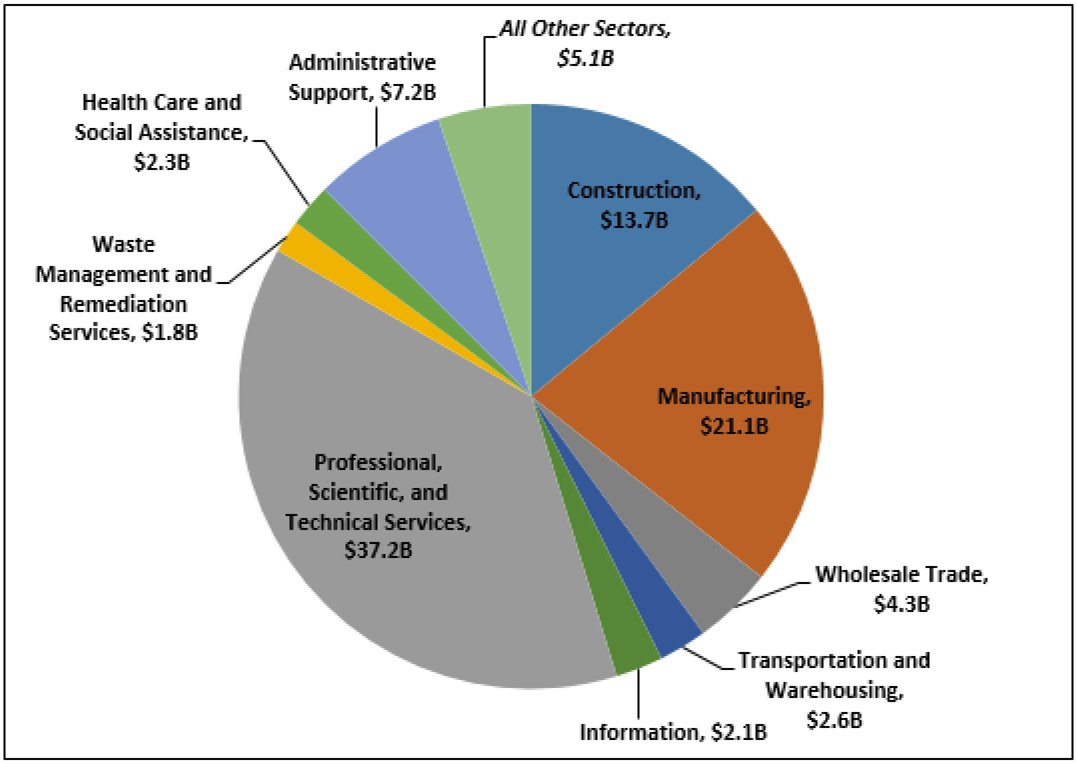


Figure 3. FY 2015 Small Business Federal Contract Obligations by Industry Sector
(Analysis of FPDS Data)

Figure 3 depicts small business federal contract obligations by industry sector. Though the same four sectors dominate, their relative sizes differ significantly. Two sectors, construction and professional, scientific, and technical services account for relatively larger pieces of the small business pie; manufacturing accounts for a noticeably smaller piece. Figure 4 compares the relative sizes of the four major sectors in each of the two procurement spaces (i.e., small business and “other than small business”).

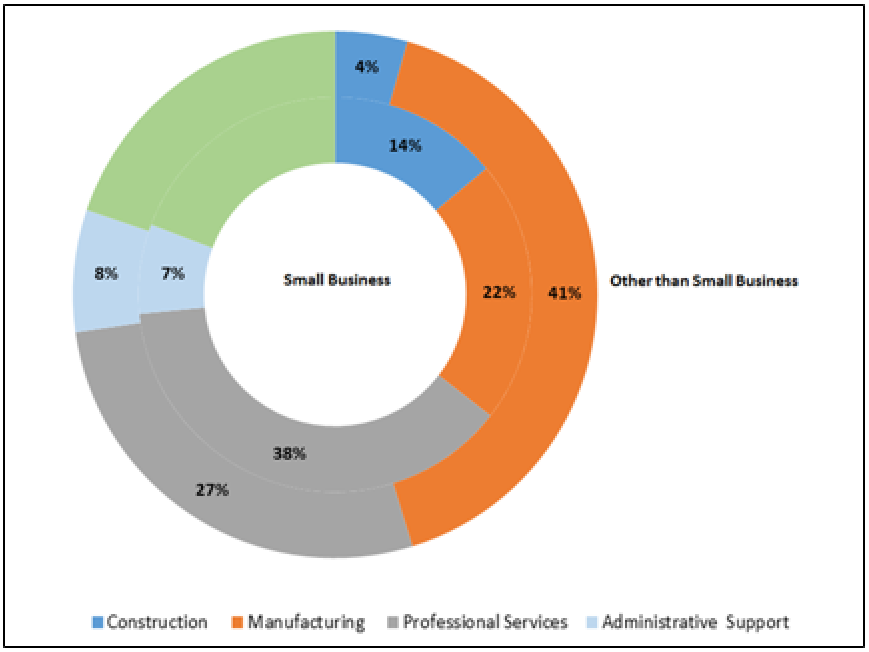


Figure 4. Sectoral Composition, Small Business vs. Other Than Small Business, Percentage of FY 2015 Action Obligations
(Analysis of FPDS Data)

The Small Business Potential

Table 3 depicts the relationship between total federal procurement within each sector and the small business share within that sector. The table shows, for example, that 4% of all federal contract obligations fall within the transportation and warehousing sector. Of that 4%, or $16 billion in total federal contract obligations, 17%, or $2.6 billion, is obligated to small business.

Increasing the small business opportunities within the 10 sectors where federal procurement is below 1% of the total will have minimal impact on the overall small business share, especially given that in eight of these "minor" sectors, small business is already well represented. In terms of federal procurement, small business dominates the agricultural sector, with 76% of all dollars (in FY 2015) awarded to small business, but this figure translates to only $318 million.

The table makes it clear that any effort to significantly increase the small business share of federal contracting dollars must be directed within the first four or five sectors, where the overall level of federal procurement is relatively high. However, there are challenges in this regard. In the construction sector, for example, nearly half (47%) of all contracting dollars already flow to small business, a figure well above SBA's government-wide small business contracting goal of 23%. As for manufacturing, the federal government spends the bulk of its contracting for manufacturing dollars in highly-specialized industries such as aerospace and military manufacturing. These industries require extensive capital investment, a large operating footprint, and far-reaching logistics networks.

Table 3. Federal Procurement by Sector and the Small Business Share, FY 2015
(Analysis of FPDS Data)

Sector	Total contract obligations	Small business share
Manufacturing	37%	13%
Professional, Scientific, and Tech Services	30%	29%
Administrative Support	8%	23%
Construction	7%	47%
Transportation and Warehousing	4%	17%
Information	3%	23%
Finance and Insurance	3%	3%
Wholesale Trade	2%	42%
Waste Management and Remediation	1%	26%
Healthcare and Social Assistance	<1%	37%
Education Services	<1%	24%
Retail Trade	<1%	37%
Other Services	<1%	27%
Real Estate and Rental and Leasing	<1%	43%
Agriculture, Forestry, Fishing, and Hunting	<1%	76%
Accommodation and Food Services	<1%	20%
Utilities	<1%	9%
Mining, Quarrying, Oil and Gas Extraction	<1%	58%
Arts, Entertainment, and Recreation	<1%	67%

Growth of Small Business in the Professional Services Sector

It seems, then, that the potential for greater small business procurement lies primarily in the professional services sectors and, to a (far) lesser extent, the administrative support and transportation and warehousing sectors. Figure 5 depicts small business trends in the four major sectors over that last decade.

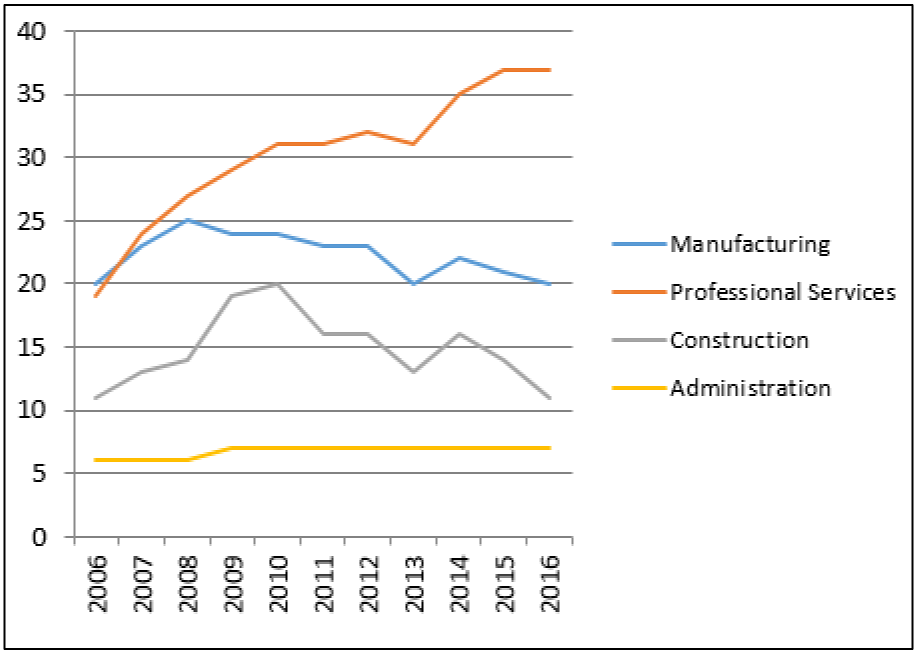


Figure 5. Small Business Share (Action Obligations in $Billions) of Federal Contract Dollars in the Four Major Sectors
(Analysis of FPDS Data)

In terms of small business representation, the graph indicates steady growth within the professional services sector. It is of note that these trends are not necessarily representative of federal procurement in general. Figure 6 shows trends in federal contracting, excluding small business, in the same four sectors.

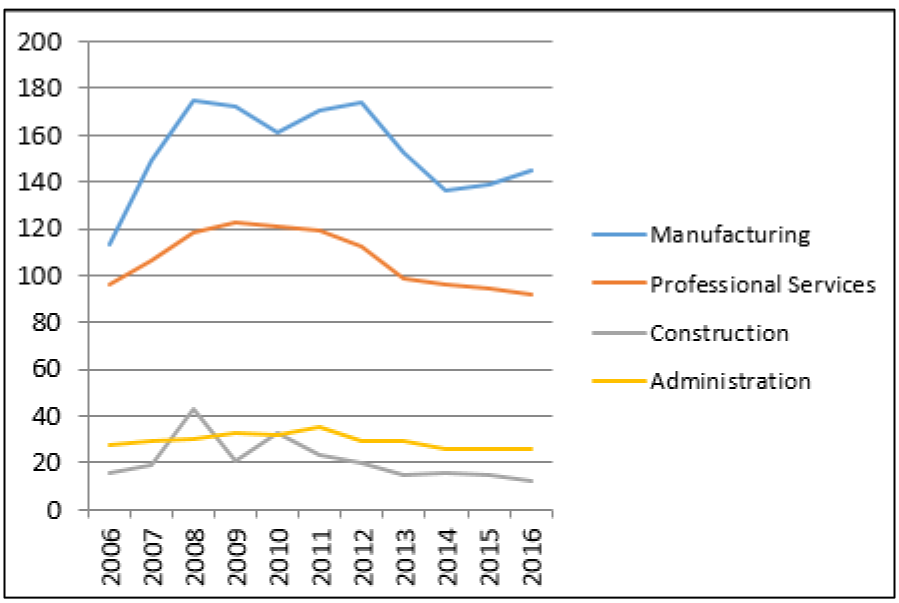


Figure 6. Federal Contracting (Action Obligations in $Billions) in the Four Major Sectors, Excluding Small Business
(Analysis of FPDS Data)

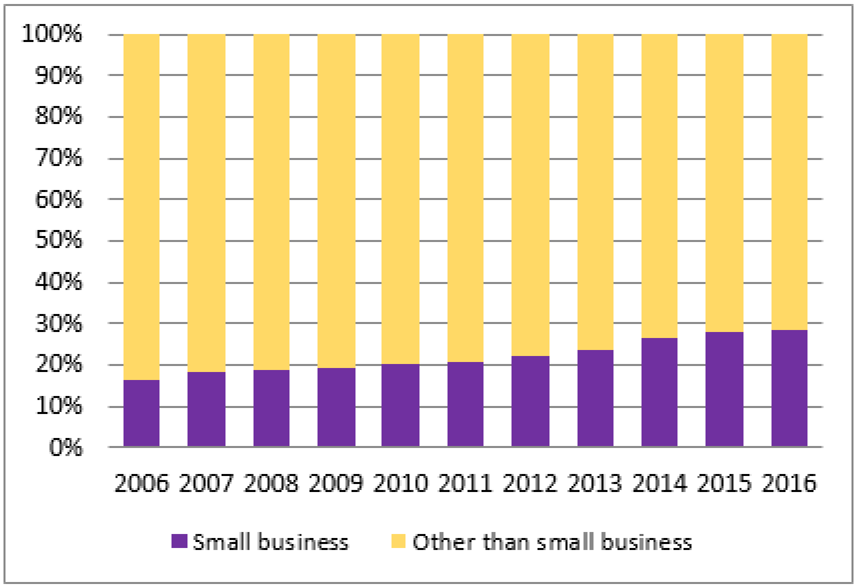


Figure 7. Small Business Participation in Federal Contracting, Professional Services Sector, Percentage of Action Obligations, $
(Analysis of FPDS Data)

It is clear that small business has lost ground in the construction and manufacturing sectors, but has gained steadily in the professional, scientific, and technical services sector. Figure 7 shows the growth of the small business share of federal contracting dollars in the professional services sector over the last decade, from 15% in 2006 to 29% in 2016.

As Table 3 indicated, within the context of federal contracting, there are very few industry sectors capable of providing significant new opportunities for small businesses. Accordingly, and as recent trends suggest, SBA set-aside policy will have the effect of concentrating more federal contract spending into the growing professional services sector.

Unintended Consequences

In this section, we highlight the unintended consequences that derive from concentrating small business contract spending into the professional services sector. For small firms, these include uneven and unsustainable growth and significant barriers to entry; for government, unintended consequences take the form of contracting and economic inefficiency.

In an effort to contextualize our findings, we present the perspectives of professional services providers (small and mid-size) as well as government officials. In both cases, their identities have been anonymized in order to solicit candid responses.

It should be noted upfront that all of the participants conveyed a favorable view of the *concept* of small business set asides. One small firm remarked that its view of set asides was

> absolutely positive. … It allows us to compete on a more level playing field. I think it's been a great program. You look at the numbers of small businesses in the United States, [and] you hear time and time again that so much of the income and GDP comes from small businesses.

Another noted that “if you didn’t have set asides, then you wouldn’t be able to seed companies.” However, when it came to the specific content of set-aside policy and its implementation, perspectives were more nuanced.

Uneven and Unsustainable Growth

Set-asides may induce the small business to grow more rapidly than it otherwise would. This growth may be uneven and unsustainable. Because the small business is not able to develop adequate depth in the provision of capabilities and other business functions in such short order, large contracts have the potential to overwhelm its infrastructure and capacity. This is an increasingly likely outcome given that small businesses are also being awarded both a higher number and greater percentage of large contracts in the professional services sector (see Figure 8), a trend that is not seen in small business procurement generally. In 2006, small business received approximately 5% of contracts over $25 million; by 2016, the figure increased to over 16%.

One mid-size business with whom we spoke offered the following perspective:

> Right now, [government agencies] are just managing against numbers. They’re managing against quotas and objectives. I think that what’s needed is a healthy step back to try to understand what is it we’re trying to accomplish. I don’t mind having small businesses get a priority for some prime contracts, but having a small business award that is a hundred million or two hundred million a year is just ludicrous. It’s totally ludicrous.

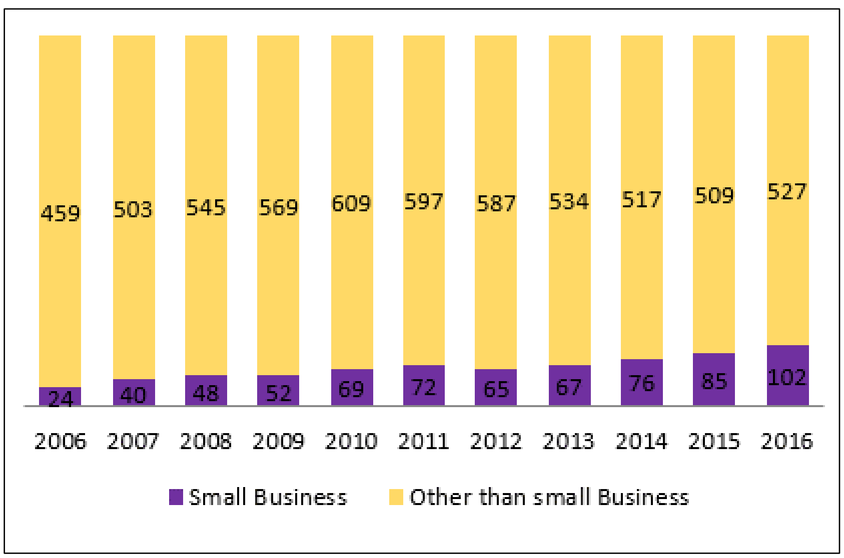


Figure 8. Number of Large Contracts (> $25 Million) Awarded Annually in the Professional Services Sector
(Analysis of FPDS Data)

As successful small businesses increase their annual revenues, graduating from certain NAICS codes, they must look to compete for government set-asides in other industries, often those whose size standards are expressed in number of employees. There are three NAICS codes in the professional services with size standards expressed in number of employees: research and development in biotechnology (1,000); research and development in the physical, engineering, and life sciences (1,000); and information technology value added resellers (1,000).

However, this transition may require that the small business reorient its business model, relinquishing the sought-after capabilities that made it successful in the first place. Clearly, this outcome represents a loss to both the firm and the government.

One small firm, whose yearly revenues recently began to exceed the 27.5 million size standard, had this to say:

> [In areas such as] software development, cybersecurity, [and] IT networking, we have to look to basically be a sub. We can't really compete for that work any longer, even [with regard to] our prime contract today … we can't re-compete for our own work there because of the NAICS code [size standards].

Indeed, this same firm decided to reorient its business model in order to pursue contracts in other NAICS codes:

> We're focusing now on other parts of our business where we do engineering services work, [which has] a higher NAICS code and research and development programs. [But] these potential prime opportunities … require a business shift for us as far as the talent that we have on board. All IT people aren't necessarily R&D people.

Another option, of course, is to compete in the full and open category alongside established mid-size firms and defense industry giants. Often, graduating small businesses are not well positioned to succeed in this environment. According to Representative Gerald Connolly (D-VA): "Innovative, high performing small businesses are becoming victims of their own success—graduating from small business programs only to find themselves in the untenable position of facing off against multi-billion dollar firms" (Weigelt, 2013).

Some small businesses may pursue yet another option: choose to limit growth and remain small to avoid disqualifying themselves for small business set-aside contracts. Rather than pursue growth and diversification so as to become independent and financially robust, they remain dependent on subsidized federal contracts to survive. These "permanent small businesses" may become quite adept in this environment over time. According to one small business executive:

> I met with another small business owner … and you won't believe this. He said "I'm in it for the nine years. I'm a retired army guy and I've also got a background so that I can be a [small disadvantaged business]. My intent is to grow it for nine years, make all the money I can and then let it die."

Needless to say, this outcome is antithetical to the SBA set-aside program's primary goal to encourage small businesses to hire more employees and grow.

Barriers to Entry

As discussed, SBA polices create market distortions by, in effect, mandating that federal agencies look increasingly to small business to fulfill their professional services requirements in order to meet SBA targets, thereby creating artificial demand for small businesses within this sector. As a result, there are more small businesses vying for a share of the overall spend, which has hindered contracting agencies' ability to effectively and efficiently contract for services. Agencies must spend more time reviewing more proposals from small businesses with which agencies may be unfamiliar. Complicating matters further is the shortage of experienced acquisition professionals. A retired senior Air Force contracting official summarized his perspective on acquisition personnel as follows:

> There is an obvious challenge when you take people who do not possess the depth of experience and you rush them into positions commensurate with

> elders who have held 15 years' worth of experience before they came into the same position. There are some obvious challenges with experience level, education and training. There are institutions out there that are trying to tackle those challenges, but textbooks and classroom training can simply not replace repetition and experience.

In an effort to circumvent the lengthy solicitation process, government agencies have turned increasingly to multiple-award indefinite delivery/indefinite quantity contracts (MA/IDIQ), often in the form of Government-wide Acquisition Contracts (e.g., the GSA's OASIS and Alliant contracts) or single-agency multiple award contracts (e.g., the Air Force's NETCENTS). Total procurement obligations under multiple-award contracts exceeded $80 billion in 2011, double the amount in 2006 (Robinson, 2013).

Another reason that the government has turned to MA/IDIQs is to avoid bid protests—i.e., a challenge to the award of a contract, typically lodged by a competitor—which have increased significantly over the last decade from 1,352 to 2,561 in 2014 (GAO, 2015). In fact, the number of annual bid protests ticked up by 5% alone in 2014, an increase that is not insignificant, considering the overall decrease in procurement spending (Burton, 2015).

Often, agencies rely on two variants of a contract, one that is reserved exclusively for small business participation and one that is "unrestricted." Small businesses that are awarded MA/IDIQ contracts compete against other small businesses for individual task orders placed by government customers. These customers often view MA/IDIQ contracts as "one stop shops" that enable them to quickly and easily meet both their professional services needs and their SBA-negotiated small business goals.

But because MA/IDIQs tend to have relatively long periods of performance, often up to five years, and few "on ramps," the contracts tend to limit participation. In essence, they create a few winners, but many other small firms are shut out. A government contract officer with whom we spoke asserted that "SBSAs are giving small businesses work, but you have to be among the select few; there are some winners but there will be a lot of losers."

According to one small defense firm,

> There are 80,000 small businesses [capable of] supporting DoD and you've got [only] 129 of them on OASIS. And the Army, Navy, and Air Force have decided all of their services work is going to OASIS. How does that support the small business industrial base? It kills it. That to me, I think, is tied to the number of protests and I think its tied to shortages in contracting officers and agencies that are so tired of dealing with all the regulations that they're looking for an easy way out.

OASIS and Contracts Consolidation

In 2013, the GSA launched its One Acquisition Solution for Integrated Services (OASIS) in response to federal agencies' requests for a more efficient process by which to hire professional services contractors. OASIS is a family of seven separate Government-wide, MA/IDIQ task order contracts spanning 28 NAICS codes and six exceptions. There are two versions of each of these contracts: one that is unrestricted and another that is reserved for small business. The contracts are referred to individually as "pools." Each of the seven small business pools is designated by a size standard. For instance, "Pool 1" consists of 21 NAICS codes that share the $15 million size standard; Pool 4 consists of two NAICS codes, both in R&D, that share the 1,000 employee size standard. When issuing an RFP under the OASIS Program, only one contract version (OASIS unrestricted or OASIS Small Business) and only one pool can be solicited.

Because OASIS has no program ceiling and a relatively long period of performance—a five-year base and one five-year option—the vehicle allows government customers to make long term plans to meet their program requirements. In small business circles, however, OASIS has been the cause of much controversy and consternation. Over the last three years, over $1.3 billion in DoD contracts alone has been transferred to OASIS, causing the abrupt displacement of numerous small businesses whose contracts are now performed by a relatively small group of 136 OASIS SB awardees.

Ironically, firms that win these coveted MA/IDIQs may not view them all that favorably because they are required to, in effect, bid twice—once for the contract and again for subsequent task orders—a process that can be onerous and expensive, especially for a small business. And, ultimately increasing overhead rates, that will be passed on to government customers.

One small firm executive categorized MA/IDIQ contracts as a "serious money drain," and stated that

> we shy away from those [MA/IDIQ contracts] tremendously. Multiple reasons. One reason is that it runs up B&P [bid and proposal] costs. You're in proposal mode constantly. Also, we've seen most of those contracts go back to the incumbents' time and time again. [And with] MA/IDIQs, there's no protest. It's not a friendly place for us to play.

MA/IDIQ contracts are not often structured to facilitate a small firm's growth. One firm with whom we spoke used to provide customized IT solutions through the Alliant GWAC. The firm noted that it had been "very successful on that contract." However, by the end of the contract's period of performance, the company had exceeded the $27.5 million size standard. According to the firm,

> We were no longer able to use that vehicle with which we were very successful. And there was no alternative. They didn't allow you on to what you might call the unrestricted, or the large business contract. They just said you're out, as if you had never won.

This firm's vice president noted that its revenues in FY 2014 and FY 2015 stood at $82 and $84 million, respectively. He stated that "this year we will close at $50; next year we will probably close below that."

To be sure, MA/IDIQ contracts can offer benefits to small business awardees that traditional contracts cannot. According to 13 CFR 121.404 "If a business is small at the time of offer for the Multiple Award Contract, it is small for each order issued against the contract." Moreover, where a concern grows to be other than small, the procuring agency may exercise options and still count the award as an award to a small business. Accordingly, a small business that exceeds the relevant size standard upon winning one or two task orders can continue to compete throughout the life of the contract, which may span five, or even ten, years.

Take, for example, Amyx, a small business founded in 1999. The firm averaged approximately $10 million per year in federal contracts between 2010 and 2013. The firm was one of the first awardees of the OASIS Small Business contract (Pool 1) when it was launched in 2013. In January of 2017, Amyx was awarded its fifth task order under Oasis valued at $189 million over five years. During the same time, Amyx was awarded other large, high-profile, contracts by the DLA. Despite having exceeded the relevant size standard, Amyx will continue to be able to compete for task orders in Pool 1 ($14 million size standard) over the course of the next seven years (Thompson, 2015).

To some, this is seen as patently unfair—as evidence that MA/IDIQs in particular, and SBA policy generally, favor a small cadre of successful firms at the expense of a much larger group that feels "shut out" from some of the federal government's most lucrative contracts. However, permitting "mid-sized" firms to compete for small business task orders under MA-IDIQs might be viewed as an apt retort to the criticism that MA/IDIQs fail to facilitate firms' growth. What is clear, however, is that the consolidation of contracts into MA/IDIQs, especially GWACs, has widened the gulf between the haves and have-nots, the winners and the losers.

Contracting and Economic Inefficiency

Critics have asserted that the timing of small business awards—concentrated at the end of the fiscal year—represents agencies' attempts to meet their annual contracting goals and/or obligate remaining agency funds (see Figure 9). By taking advantage of set-aside policy to bypass lengthy sourcing, agencies are able to obligate their remaining funds quickly. One small business with whom we spoke provided the following perspective:

> You know that the fiscal year ends for the government in September 30. You also know that the federal government is not a business where they earn profit or a return on investment—their goal is to spend all of their money as fast as they can, so that they can continue to get the same level of funding. So when they get to around the August timeframe, they realize how much money they have left. If there are some things that they are interested in and a small business is able to bring that value to them, they can quickly put a sole source out and get rid of that money.

Needless to say, efforts made to spend funds quickly likely fail to maximize taxpayer value, representing yet another unintended outcome of set-aside policy.

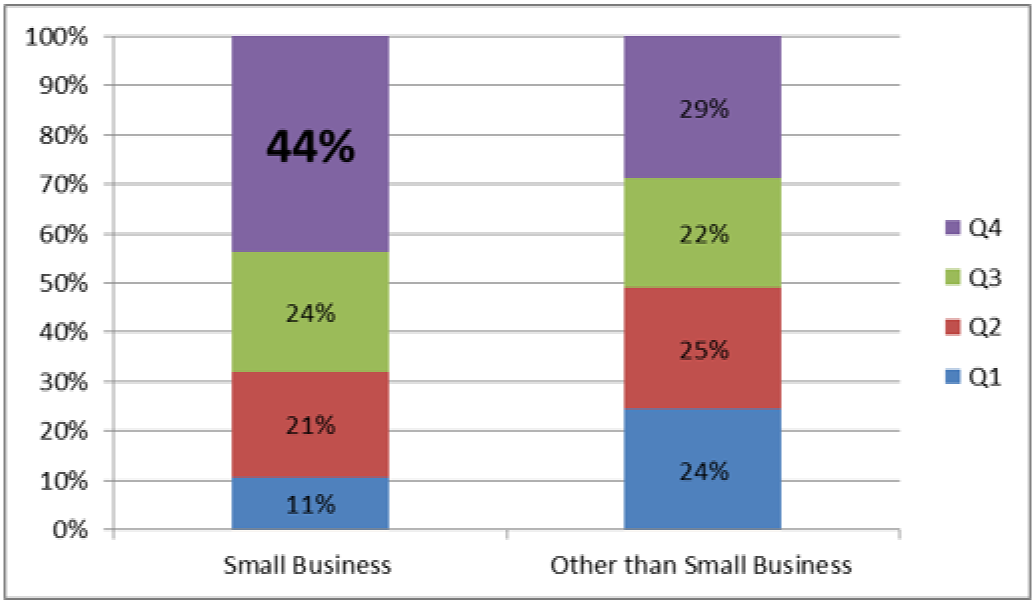


Figure 9. Timing of Contract Awards ($), 10-Year Average, 2007–2016
(Analysis of FPDS Data)

NAICS code selection can be another source of inefficiency. With the increasing pressure to meet the small business set-aside goals, agencies, on occasion, use an inappropriate NAICS code. In fact, the term "code shopping" has emerged to describe agencies' attempts to use NAICS codes with larger size standards, ostensibly in an effort to attract larger small business, with enough staff to meet user requirements, yet still meet their small business contracting goals. Because NAICS codes within the professional services sector tend to be more open to subjective interpretation than codes in other sectors, and because professional services firms often provide diverse and varied services under a single contract, there is greater potential for code shopping within the growing professional services sector.

Set-aside policy also creates the potential for significant economic inefficiency within the professional services sector. As one mid-sized business executive observes,

> The government is always prone, when it hears about any inequities, to create more categories, more numbers, more demographic barriers, or segments. We continue to see the proliferation of size standards and demographic categories. At some point you have to ask, does the creation of these categories become counterproductive? By segmenting the industry space, do you force turbulence? Do you force unnecessary churn in the market?

With the current set-aside policy, mid-size and large professional have fewer opportunities. A Center for Strategic and International Studies (CSIS) study concluded that mid-size firms (these were defined as firms which are too large to be categorized as small but had less than $3 billion in total annual revenue) were being "squeezed" out of DoD contracts by both large and small contractors. CSIS found that from 1999 to 2009, the share of DoD contracts awarded to small businesses increased (from 17.0% to 17.4%) and to large firms increased (from 47.0% to 53.7%), while the share awarded to these mid-size firms decreased (from 36.0% to 28.9%; CSIS 2012).

Regarding the concentration of small firms in the professional services sector, a mid-sized defense firm with whom we spoke offered the following perspective:

> It doesn't make sense to have all of our services business go to small business, because quite candidly, I'm not sure what that really does for the nation. To all of a sudden have these body shops that are now small businesses … well, these businesses often struggle.

The firm commented that its defense customers began to turn to small businesses in 2012 and 2013, when the government placed renewed emphasis on meeting set-aside goals. The firm noted that

> as our contracts came up for re-compete, our customers were very up front about it. They said, "Hey, we don't want to go small business, we don't think it makes any sense. But we are being forced to go small business." So we saw a very significant squeeze. A contract may have had 20 or 30 of our people and now it's up for re-compete; all of a sudden, it's going to be a small business contract.

The intent of the small business set-aside programs is to grow small, and, in many cases, disadvantaged businesses to become competitors for additional contracts. As the awards grow in size, the small firms are forced to team with a larger firm as a sub-contractor. When these companies exist simply as shells or as "pass-throughs," they fail to meet the objectives of the SBSA program. According to a senior defense official:

> Anytime the small business is working in name only, this causes the DoD to simply pay a mark-up fee of 2% to 8%. This is detrimental and unfair to the taxpayer when we blindly give work to smalls, and this happens a lot because the government focuses on excessive amount of small businesses, that's when you get the shell companies to emerge.

Findings and Recommendations

Small business procurement is not evenly distributed across industry sectors; rather, it is increasingly concentrated in a few sectors, including, most notably, the professional services sectors. Consequently, small business set-aside policy, in its current implementation, yields negative consequences for small business and for government.

Findings

- Often, agencies' small business contracting goals must be met by "overspending" in the professional services sector, which creates artificial demand for small professional services providers, while "squeezing out" established, and often better qualified, mid-sized and large firms.
- SBA policies have clearly facilitated the growth in the award going to small businesses, however uneven or, at times, unsustainable. However, the rationale or methodology for developing the goals, if one exists, is not well understood. Additionally, since current revenues form the basis of future size standard determinations, many growing industries within the professional services sector will be subjected to upward revisions, thereby raising the barriers to new entrants, perhaps hindering innovation.
- When small professional services providers receive larger contracts, their growth trajectory is accelerated, such that they are often no longer eligible for set-asides; these providers frequently lack the capability and infrastructure to compete under free and open competition.

- Agency set-aside goals encourage a trend whereby small business must subcontract with a large business that will perform work in areas where the small business has limited capabilities. In some instances, the small business acts as “a pass-through” that offers limited or even negative value to government.
- The proliferation of small firms in the professional services sector, combined with a declining acquisition workforce, has fueled increased reliance on multiple award contracts, which favor a select group of small firms, but “shut out” the majority.
- In an effort to meet small business goals, agencies may resort to “code shopping” in an effort to obtain the best of both worlds: the services of a larger, more qualified, “small” business and credit towards their small business contracting goals.
- When government agencies need to obligate funding quickly, they turn to small business contracting, a practice that may not obtain government best value.
- The complex regulatory environment, especially within the DoD, SBA size standards (revenue or number of employees) for small business that vary across more than one thousand industries, in addition to goals for prime and subcontracting that differ by agency and type of small business (e.g., minority-owned, women-owned, etc.) all require a large bureaucracy that is maintained at taxpayer expense.

Based on our analysis of FPDS data and our examination of the unintended consequences that derive from set-aside policy, we offer the following recommendations.

Recommendations

Set Realistic Numeric Agency Goals

The Small Business Administration should develop an understandable, rational, and transparent, methodology to establish numeric agency small business goals. Consideration should be given to the development of a single goal, that would include both prime and subcontract dollars; adjusting the calculation and grading methodology to account for small business firms participating at all tiers of the agency contracts.

Encourage the Best Small Businesses to Grow

Small business set-aside program must be structured to encourage the best firms to grow. To accomplish this, there must be enough opportunities for these firms to compete when they graduate.

Improve Data Gathering and Program Metrics

It is impossible to understand the full impact of any program without reliable data and metrics. Currently, the small business set-aside program focuses on achieving the numerical goals for small business contracting. However, it is difficult to assess the costs that this program may impose, and how successful the program is in achieving the program’s overall objectives (e.g., job creation, innovation, growth, etc.), that is, the program’s outcomes.

Use Set-Asides for Acquisitions Only When Small Business Can Handle Them

When given the appropriate contracts, small businesses can successfully perform as, or more, efficiently than a large business. The key is selecting the suitable opportunities that are within the scope and scale of the small business, so that selecting a small business prime does not create a risk of poor performance. Small business set-asides are suitable when they enable a firm to grow, but do not overwhelm its infrastructure or capabilities. Awarding a contract to a firm that is beyond its capacity will cause the company to have difficulty with that work, and may cause it to fail. Agencies should refrain from awarding large contracts that approach or exceed the industry size standard. Large contracts have the potential to overwhelm small firms' infrastructure and capabilities. Moreover, these contracts prematurely hasten a small firm's growth trajectory, often to point where the firm is no longer eligible to receive set-asides.

Review NAICS Code Thresholds

The SBA has defined these size standards for groups of industry. When these groups are too broad, the codes can provide enough ambiguity so that an inappropriate code (and as a result size standards) can be used. This results in an inappropriate set of firms that are subsidized, and the intended recipients are not eligible. These thresholds must be clear and unambiguous.

Review the Use of Multiple Award/IDIQ Contracts

Reliance on IDIQ vehicles as convenient tools for flexible contracting has helped reduce the transaction costs associated with many programs. However, IDIQ contracts have the potential to limit overall competition since potential vendors are preselected. Small businesses that are not awarded IDIQs in their industry are effectively "shut out" of some of the federal government's most lucrative contracts. Those fortunate enough to be awarded IDQs face high bid and proposal costs (relative to traditional contract solicitations) in that they must bid on the initial contract and then again for each individual task order placed under that contract). For small businesses with limited means, these costs may prove prohibitive, creating barriers to entry and constraining innovation.

Conclusion

As Milton Friedman once remarked, "One of the great mistakes is to judge policies and programs by their intentions rather than their results." Current federal policy with respect to small business set-asides was formed, and is implemented, with the best of intentions. However, as with many policy initiatives, there can be unintended consequences. The government must strike a balance that encourages the growth of innovative small businesses while ensuring that its contracting needs are met in a way that is responsible, effective, and efficient. Small business set-aside policy, in its current implementation, may not strike the optimal balance.

References

Burton, R. (2015). Annual review of government contracting. *Bloomberg Government*. Retrieved from http://www.ncmahq.org/docs/default-source/default-document-library/pdfs/exec15---ncma-annual-review-of-government-contracting-2015-edition

CSIS. (2012). *Defense contract trends: U.S. Department of Defense contract spending and the supporting industrial base*. Washington, DC: Author.

FAR 19.201. Small business programs, general policy. Retrieved from https://www.acquisition.gov/far/html/FARTOCP19.html

FAR 19.502. Setting aside acquisitions. Retrieved from https://www.acquisition.gov/far/html/Subpart%2019_5.html

GAO. (2015). *GAO bid protest annual report to Congress for Fiscal Year 2015*. Washington, DC: Author. Retrieved from http://www.gao.gov/assets/680/674134.pdf

McVay, B. (2009). Set-asides and preference programs. In *Getting started in federal contracting.* Retrieved from http://www.fedgovcontracts.com/chap11.pdf

Obama, B. (2014, May). *National Small Business Week, 2014* [Presidential proclamation]. Washington, DC: The White House, Office of the Press Secretary. Retrieved from https://obamawhitehouse.archives.gov/the-press-office/2014/05/09/presidential-proclamation-national-small-business-week-2014

Robinson, B. (2013, July). Unraveling the facts about GWACs. *FCW*. Retrieved from https://fcw.com/articles/2013/06/07/feature-gwac-facts.aspx

SBA. (2014). Mission. Retrieved from http://www.sba.gov/about-sba/what_we_do/mission .

SBA. (2015). Qualifying as a small business. Retrieved from https://www.sba.gov/contracting/getting-started-contractor/qualifying-small-business

SBA. (2016). Goaling. Retrieved from https://www.sba.gov/contracting/contracting-officials/goaling

Small Business Act of 1953, 15 U.S.C. § 631 (1953).

Small Business & Entrepreneurship Council. (2016, October). Lost manufacturing, lost small manufacturers. Retrieved from http://sbecouncil.org/2016/10/18/lost-manufacturing-lost-small-manufacturers/

Thompson, M. (2015). U.S. Air Force awards Amyx task order on OASIS small business pool 1 contract for cyber protection team. *Washington Executive*. Retrieved from https://www.washingtonexec.com/2015/11/u-s-air-force-awards-amyx-task-order-oasis-small-business-pool-1-contract-cyber-protection-team/

Weigelt, M. (2013, February 19). Set-aside for mid-tier companies still possible. *FCW*. Retrieved from http://fcw.com/articles/2013/02/19/mid-tier-set-aside.aspx

Acknowledgment

This research was partially sponsored by a grant from the Naval Postgraduate School.

Exploring Drivers of Better Strategic Sourcing in the Air Force Using Analytics

Aruna U. Apte—is an Associate Professor in the Operations and Logistics Management Department of the Graduate School of Business and Public Policy at the Naval Postgraduate School in Monterey, CA. Aruna has successfully completed various research projects, involving applications of mathematical models and optimization techniques that have led to over 20 peer-reviewed journal articles, over 40 research articles, and one patent. Her research interests are in developing mathematical models for complex, real-world operational problems using optimization tools. She values that her research be applicable. Currently her research is focused in humanitarian and military logistics. She has several publications in journals, such as *Interfaces*, *Naval Research Logistics*, *Production and Operations Management*, *Decision Sciences Journal*. She has also published a monograph on Humanitarian Logistics.

Aruna received her PhD in Operations Research from Southern Methodist University in Dallas. She also has a M. A. in Mathematics, from Temple University, Philadelphia. Before NPS she worked as a consultant at MCI and taught at the Cox School of Business, SMU, where she won the best teacher award. She has over 20 years of experience teaching operations management, operations research, and mathematics courses at the undergraduate and graduate levels. At NPS She teaches mathematical modeling, for which she won the best teacher awards (Liskin and Hamming) and has advised over 99 students for 46 MBA/Masters reports. She has also advised emergency planners in preparing for disaster response. She is the founding and past president for a college (focus group) in Humanitarian Operations and Crisis Management under the flagship of academic professional society in her intellectual area of study, Production and Operations Management Society. [auapte@nps.edu]

Karen A. F. Landale—is an Assistant Professor at the Graduate School of Business and Public Policy, Naval Postgraduate School, Monterey, CA. She teaches strategic sourcing and category management as part of the contracting curriculum. Major Landale received her PhD from the Kenan-Flagler Business School at the University of North Carolina at Chapel Hill. Her research focuses on talent management, services marketing, and contracting in the public domain. She is married to LTC Gordon Landale, USA, and they have one daughter, Amelia. [kalandal@nps.edu]

Rene G. Rendon—is an Associate Professor at the Graduate School of Business and Public Policy, NPS, where he teaches defense acquisition and contract management courses. He also serves as the Academic Associate for the MBA specialization in contract management. Prior to joining the NPS faculty, he served for over 20 years as an acquisition contracting officer in the United States Air Force. His career included assignments as a contracting officer for the Peacekeeper ICBM, Maverick Missile, and the F-22 Raptor. He was also a contracting squadron commander and the director of contracting for the Space Based Infrared Satellite program and the Evolved Expendable Launch Vehicle rocket program. Rene has published in the Journal of Public Procurement, the *Journal of Contract Management*, the *Journal of Purchasing and Supply Management*, and the *International Journal of Procurement Management.* [rgrendon@nps.edu]

Javier Salmerón—is an Associate Professor in the Operations Research Department at the Naval Postgraduate School. His research focuses in the area of mathematical modeling and optimization, and its applications in civilian and military problems. This has been published in open, peer-reviewed journals. He also teaches several graduate courses in optimization and operations management, and advises thesis students. Before his time with NPS, Dr. Salmerón worked for Spanish electric utility Iberdrola. He was also part-time adjunct professor in the Department of Statistics and Operations Research at the Statistics School of Complutense University of Madrid. In 2009, he and other department colleagues received the American Institute of Aeronautics and Astronautics Homeland Security Award, and in 2014 he and other department colleagues and students received the Military of Operations Research Society Richard H. Barchi Prize. In 2015 he and K. Wood received the Best Paper Award at the 5th International Defense and Homeland Security Simulation Workshop. He also received the Wayne E. Meyer Award for teaching excellence in Systems Engineering in 2012 and 2014. He holds SCI clearance. [jsalmero@nps.edu]

Introduction

The U.S. Department of Defense (DoD) annually obligates billions of dollars for the procurement of supplies and services in support of the national military strategy. In fiscal year (FY) 2016, the DoD obligated approximately $239 billion on contracts for defense-related supplies and services. Specific to the U.S. Air Force (USAF), over $50 billion were obligated on contracts for supplies and services in FY2016 (USA Spending, 2016). Services typically account for over half of the DoD procurement budget, compared to the acquisition of supplies. In the current environment of budget and manpower cuts, the DoD is transforming its acquisition process to ensure that critical supplies and services are sourced cost-effectively.

The DoD has been undergoing a transformation of its procurement function from a transaction-oriented perspective to a strategic-oriented enterprise. The procurement function is no longer seen as a tactical, clerical, or administrative function, but more of a strategic function. This transformation can be attributed to the fact that the DoD has begun to understand and realize the importance of procurement in achieving the strategic objectives as well as the impact of procurement on reducing costs. One aspect of this transformation in the DoD is the use of a strategic sourcing approach, specifically category management, for the procurement of services at military installations. Category management is a federal government initiative that emphasizes a focus on “increased efficiency and effectiveness, lessening costs, and reducing redundancies” (Sharkey, 2015). Category management emphasizes leveraging buying power, improving efficiencies, and managing consumption. The Air Force is leading the DoD in its category management initiative through category planning, category execution, and category performance management. Specific to category execution, the use of performance levers such as total cost management and the identification of specific cost drivers in service acquisitions can result in increased efficiency and effectiveness and a reduction in costs.

Purpose of Research

The purpose of this research is to analyze the price drivers for one of the DoD’s most commonly procured installation-level services, integrated solid waste management (ISWM). Specifically, we focus on the procurement of ISWM services within the Air Force to identify the relationship between service-related price drivers, contract-related price drivers, price, and contractor performance. Our focus is to study the effect that price drivers (both service and contract) have on contract price and contractor performance. We test seven hypotheses to determine the effect that service and contracting variables have on price and contractor performance. Based on our research findings, we provide recommendations to the Air Force for strategically sourcing ISWM services that will result in increased efficiency, effectiveness, and a reduction in costs.

Our Previous Work

In our previous research on the Air Force strategic sourcing process, we developed an optimization model for selecting a set of proposals from among multiple offerors for services to be performed at multiple installations (Apte, Rendon, & Salmerón, 2011). The selection achieved the most favorable objective by balancing the confidence level in an offeror’s past performance with the cost of services to the Air Force. The research findings, which were based on a realistic scenario, demonstrated improvements over the sourcing process in both overall performance and cost.

We continue our research stream with this current research methodology using analytics, specifically statistical analysis, to explore price drivers for optimal strategic sourcing of ISWM services. We focus on price as the principal driver and use performance data to understand the correlation between pricing and performance of the contractors. Our basis for this research is that insight into pricing and performance will help strategize the sourcing of contracts for the decision-makers. We use three statistical methods to determine how service-related and contract-related independent variables (tonnage of waste, number of containers, wage rates, number of offers, and type of completion) affect the dependent variables (total price and contractor performance).

Literature Review

Procurement Transformation

The transformation of the procurement function from a transaction-oriented perspective to a strategic-oriented focus was first discussed by Henderson (1975, p. 44) when he predicted that there would be greater importance placed on the procurement function in corporate management. Kraljic (1983) purported that "purchasing must become supply management" and that organizations should develop specific sourcing strategies for products/services based on the strategic importance of the procured supply/service to the organization and the complexity of the market for that product/service. Kraljic developed a systematic framework for incorporating environmental and other strategic factors into procurement strategy formulation for procured products/services. The use of the Kraljic framework results in a contingency-based model for formulating the appropriate sourcing strategy for products/services. The Kraljic framework has been widely applied throughout the industry. Rendon and Templin (1992) explored the application of the Kraljic framework to National Cash Register's (NCR's) supply management program. The use of the Kraljic framework enables the organization to determine the appropriate sourcing strategy for specific products/services. The market complexity and importance of the product/service to the organization may indicate that a strategic sourcing strategy is appropriate.

Strategic Sourcing and Category Management

One aspect of the purchasing transformation to a strategic function is the use of a strategic sourcing approach for the procurement of product/services. Strategic sourcing involves taking a strategic approach to the selection of suppliers—an approach that is more aligned with the organization's strategic objectives and reflects the integration of sourcing with corporate strategy (Rendon, 2005, p. 9). Closely related to strategic sourcing is category management, which is concerned with the strategic sourcing of a specific category of product/services to ensure the sourcing of those products/services meet corporate-level strategic objectives. Strategic sourcing is differentiated from category management in that strategic sourcing is a one-time event that is focused primarily on leveraging to drive down costs. Category management is an ongoing process that is focused on value elements that go beyond simple price savings. Category management involves engaging stakeholders and fully understanding their product/service requirements, market intelligence on market trends, cost drivers, and risks pertaining to those product/services, and developing a sourcing strategy that aligns stakeholder requirements with the realities of the market (Monczka et al., 2015, pp. 199–201).

Federal Government and Air Force Initiatives

The federal government has implemented both strategic sourcing and category management as part of its initiatives to reduce costs and increase efficiency and effectiveness. The U.S. Government Accountability Office (GAO) identifies five principles of strategic sourcing: maintaining spend visibility, centralizing procurement, developing category strategies, focusing on total cost of ownership, and regularly reviewing strategies and tactics (GAO, 2016). As reported by the Office of Management and Budget (OMB), federal agencies have

> saved money by pooling their spending, either by centralizing the agency's contracting decisions or by using government-wide strategic sourcing vehicles, in order to lower prices and reduce duplication and administrative costs. Since FY 2010, government-wide contracts for office supplies have saved over $140 million by offering lower prices than any single agency could negotiate on its own. Similar vehicles for domestic delivery services saved over $31 million in fiscal year (FY) 2011 over what agencies were paying under previous agreements. (OMB, 2012, p. 1)

Through its initiatives such as "Buying as One Through Category Management," the OMB is focusing on "managing commonly purchased goods and services … by implementing strategies to drive performance, like developing common standards in practices and contracts, driving greater transparency in acquisition performance, improving data analysis, and more frequently using private sector (as well as government) best practices" (OMB, 2014, p. 2).

The U.S. Air Force is leading the DoD in its category management initiative by focusing on strategic sourcing savings levers of "leveraging buying power, improving efficiencies, and managing consumption" (Sharkey, 2015, p. 7). The Air Force's category management operating model includes category planning, category execution, and category performance management.

Category planning involves conducting a spend analysis, requirement analysis, market analysis, and risk analysis. This phase also includes analyzing the four major performance levers (demand management, supplier management, strategic sourcing, and total cost management) to identify category improvement initiatives. Category improvement initiatives within total cost management includes the identification of specific price drivers in the acquisition that can result in increased efficiency and effectiveness and a reduction in costs. Price drivers can be either product/service-related or contract-related and impact savings associated with rate (getting more for less), process (getting more with less), and demand (getting less) (Sharkey, 2015, pp. 21–24). The product/service-related price drivers impact rate savings, process savings, and demand savings. Contracting-related price drivers impact rate savings.

Category execution involves the execution of selected performance levers identified in the planning phase. This would include executing changes associated with the product/service-related or contract-related price drivers (Sharkey, 2015, pp. 25–30).

Category performance management includes the performance tracking, benchmarking, and continuous improvement of the management of the specific category of product/service (Sharkey, 2015, pp. 31–33).

Academic Research

As previously stated, the purpose of this research is to analyze the price drivers for the Air Force's procurement of integrated solid waste management to identify the relationship between service-related price-drivers, contract-related price drivers, price, and contractor performance. Our focus is to study the effect that price drivers (both product/service-related and contract-related) have on contract price and contractor performance. Our research fills a gap in the ISWM literature. Past research has focused mostly on solving waste management social and environmental problems. For example, Achillas et al. (2013) conducted a literature review on multi-criteria decisions aiding in waste management problems for all reported waste streams. Their review provides decision-makers with a thorough list of practical applications of the multi-criteria decision analysis techniques that are used to solve real-life waste management problems.

The waste management literature also includes research exploring the most cost-effective waste collection system. For example, Boskovic et al. (2016) developed a management tool to determine waste collection costs for different waste collection schemes and input data (waste quantity and composition, the number of waste bins, the location of collection points, the type of collection vehicle, crew, and collection route). The tool can calculate the time and costs of waste collection (per vehicle, collection point, or ton of collected waste).

Additionally, Arribas, Blazquez, and Lamas (2010) conducted case study research which proposed a methodology for designing an urban solid waste collection system which uses combinatorial optimization and integer programming and geographic information system tools to minimize collection time and operational and transport costs while enhancing the current solid waste collection system. Their methodology establishes feasible collection routes, determines an adequate vehicle fleet size, and presents a comparative cost and sensitivity analysis of the results. Their research findings yielded significant cost savings in the total solid waste collection system.

Finally, Solano et al. (2002) developed an integrated solid waste management model to assist in identifying alternative solid waste management strategies that meet cost, energy, and environmental emissions objectives. They categorize waste into 48 items and their generation rates were defined for three types of sectors: single-family dwelling, multifamily dwelling, and commercial. The model is flexible to allow representation of waste diversion targets, mass flow restrictions and requirements, and targets for the values of cost, energy, and emission.

As previously stated, the purpose of this research is to analyze the price drivers for the USAF procurement of integrated solid waste management services to identify the relationship between service-related price-drivers, contract-related price drivers, price, and contractor performance. The next section is a discussion of our research methodology.

Methodology

Data

We used data from USAF contracts for ISWM across 63 bases. These data were originally collected by the Facilities and Construction Category Management Team, Facility Related Services subcategory. The team's goal in collecting the data was to better understand the ISWM needs across all pertinent bases in order to strategically source the service. Specifically, the team was looking for potential rate (i.e., price), process (i.e., ordering and delivery of the service), and demand (i.e., ordering the right amount of the service to meet needs) savings.

In this study, we use the data to determine the relative importance of each of the ISWM price drivers. Price of the contract is from the viewpoint of the customer, USAF, whereas the cost of providing the service is from the viewpoint of the vendor. Further, we examine the effect of small business set-asides on the price of the service by comparing price premiums of contracts that used one of five different small business set-aside categories to the price paid for contracts that used full and open competition (i.e., no small business set-asides).

The data pertinent to this study consist of 17 variables and 57 samples. Variable descriptions and types (dependent variable, DV, or independent variable, IV) are given in Table 1. For performance DVs, the buyer rates the contractor's performance on a 1-to-5 scale, where: 1 = Unsatisfactory; 2 = Marginal; 3 = Satisfactory; 4 = Very Good; and, 5 =

Exceptional. Basic descriptive statistics are given in Table 2. Correlations are provided in Table 3.

Hypotheses

Using these data, we test seven hypotheses. We begin with the price-related hypotheses, then move to the performance-related hypotheses.

Table 1. List of DVs and IVs Used in the Study

Variable name	Description (including units or rating scale)	Type
Total Price	Total price of the ISWM contract ($)	DV
Contractor Performance—Quality	Buyer-rated assessment of the contractor's performance related to quality (1-5)	DV
Contractor Performance—Cost	Buyer-rated assessment of the contractor's performance related to cost (1-5)	DV
Contractor Performance—Schedule	Buyer-rated assessment of the contractor's performance related to schedule (1-5)	DV
Contractor Performance—Small Business Subcontracting	Buyer-rated assessment of the contractor's performance related to meeting small business subcontracting requirements (1-5)	DV
Contractor Performance—Management	Buyer-rated assessment of the contractor's performance related to management (1-5)	DV
Contractor Performance—Average Rating	Average of all available performance ratings (quality, cost, schedule, small business subcontracting, management) (1-5)	DV
Tons of Waste	Annual amount of solid waste (tons)	IV
Number of Containers	Number of dumpsters serviced by the ISWM contract (dumpsters)	IV
Wage Rate	Dollars per hour paid to ISWM contractors ($/h)*	IV
Number of Offers	Number of offers received and evaluated prior to contract award (offers)	IV
8(a) Sole Source—SB Set-Aside**	Contract was provided without competition to a qualified 8(a) contractor (yes or no)	IV
8(a) Competed—SB Set-Aside**	Contract was competed among qualified 8(a) contractors (yes or no)	IV
HUBZone—SB Set-Aside	Contract was competed among qualified HUBZone contractors (yes or no)	IV
Service-Disabled Veteran-Owned—SB Set-Aside	Contract was competed among qualified SDVOSB contractors (yes or no)	IV
Total Small Business Set-Aside	Contract was competed among all qualified small businesses (yes or no)	IV
Full & Open Competition	Contract was competed among all qualified contractors (large and small) (yes or no)	IV

* Wage rate was determined using the Department of Labor rates required for all federal contracts.
** 8(a) (named after legislation that created the program) is for a special category of disadvantaged, small businesses that require significant development. The program assists those companies by offering special set-asides and even sole-source awards.

Table 2. Basic Descriptive Statistics

Variable name	Number of observations	Mean	Standard deviation	Minimum	Maximum
Total Price ($)	57	255,321.60	219,698.90	9,234	1,091,814
Contractor Performance—Quality	32	4.31	.69	3.00	5.00
Contractor Performance—Cost	10	4.00	.82	3.00	5.00
Contractor Performance—Schedule	32	4.28	.63	3.00	5.00
Contractor Performance—Small Business Subcontracting	5	3.60	.89	3.00	5.00
Contractor Performance—Management	32	4.31	.82	2.00	5.00
Contractor Performance—Average Rating	32	4.28	.63	3.00	5.00
Tons of Waste (tons)	54	1,868.77	1,937.73	75.00	10,320.00
Number of Containers	50	124.06	103.46	8.00	494.00
Wage Rate ($/h)	57	14.95	3.01	9.72	22.92
Number of Offers	50	3.24	2.54	1.00	10.00
8(a) Sole Source—SB Set-Aside	57	.19	.40	0 (no)	1 (yes)
8(a) Competed—SB Set-Aside	57	.09	.29	0 (no)	1 (yes)
HUBZone—SB Set-Aside	57	.05	.23	0 (no)	1 (yes)
Service-Disabled Veteran-Owned—SB Set-Aside	57	.02	.13	0 (no)	1 (yes)
Total Small Business Set-Aside	57	.28	.45	0 (no)	1 (yes)
Full & Open Competition	57	.25	.43	0 (no)	1 (yes)

Table 3. Correlations

Correlations									
	Total Price	CP Quality	CP Cost	CP Schedule	CP Mgmt.	Tons	Number of Containers	Wage Rate	Number of Offers
Total Price	1.00								
CP Quality	.46	1.00							
CP Cost	.30	.67	1.00						
CP Schedule	.03	.71	.42	1.00					
CP Mgmt.	.53	.91	.79	.42	1.00				
Tons	.46	.39	.01	-.22	.43	1.00			
Number of Containers	.72	-.08	-.06	-.17	.01	-.10	1.00		
Wage Rate	-.49	.06	.35	.11	.22	-.46	-.26	1.00	
Number of Offers	.03	.36	.26	.53	.26	-.24	-.20	-.01	1.00

*Note: "Contractor Performance—Small Business Subcontracting" is not included due to the small number of observations.

Price-Related Hypotheses

The first hypothesis seeks to determine the relative importance of each of the price drivers of the ISWM service. Knowing the price drivers is important in determining how the organization can control, and if possible, reduce price. We are interested in understanding whether ISWM service-related variables or contracting-related variables contribute, and if so, identifying the largest price drivers. Service-related price drivers may be able to be controlled or reduced by changing certain organizational activities. Similarly, identifying significant contracting-related price drivers can help the organization craft better acquisition strategies to control or reduce overall price. We test that the ISWM service-related variables will have more effect on the price than the contracting-related variables. Specifically, we hypothesize that the tonnage of waste, number of containers, and wage rate will influence price more than the number of offers received or the type of small business set-aside (if any).

H1: ISWM service-related variables have a greater effect on price than contracting-related variables.

The second hypothesis tests the relative effects the service-related variables have on price. While both tonnage of waste and the number of containers to be emptied logically contribute to overall price of the contract, we speculate that tons of waste has a greater effect on price because more tonnage requires more contracted trucks to dispose of the waste, and it also increases landfill costs (assuming the landfills have either a “per truck” or “per ton” fee). Further, because federal contractors are required to use standard Department of Labor wage rates (in dollars per hour) when estimating their costs, we test that wage rates will have less effect on the overall price of the service (because the wage rates are pre-determined).

H2a: Tonnage of waste has a greater effect on price than number of containers.

H2b: Number of containers has a greater effect on price than wage rate.

The third hypothesis tests the relative effects the contracting-related variables have on price. Again, while both small business set-asides and the number of offers received logically affect overall price of the contract, we posit that limiting competition through the use of set-asides has a greater effect on price because, unlike large businesses, small businesses typically do not have the volume of work required to offer deep discounts. Therefore, even if a small business set-aside contract were to receive the same (or more) offers than a full and open competition contract (i.e., a contract that allows any business to compete, regardless of size), the prices offered by small businesses are likely to be higher than prices offered by large businesses.

H3: Small business set-asides have a greater effect on price than number of offers.

The fourth hypothesis tests the effect small business set-asides have on price. As described above, small businesses typically cannot match or beat the prices of larger businesses. We use two standardized price variables to examine the effect of small business set-asides on price: (1) price per ton, and (2) price per container.

H4a: Small business set-asides result in a higher price per ton than full and open competition.

H4b: Small business set-asides result in a higher price per container than full and open competition.

Finally, the fifth hypothesis tests the relative effect the small business set-aside categories have on price per ton. There are five different small business set-aside categories represented in the data: 8(a) Sole Source, 8(a) Competed, HUBZone, Service-Disabled Veteran-Owned Small Business (SDVOSB), and Total Small Business set-aside. Of these five categories, the first four are less inclusive than the fifth. The category “Total Small Business” allows for any small business to compete for the contract—including any businesses that are in the first four categories—however, the reverse is not true. For example, if the contracting officer were to specify that the contract is a Total Small Business set-aside, any small business category is able to compete for the contract. However, if the contracting officer were to specify that the contract is a HUBZone set-aside, only those small businesses that qualify for HUBZone status are eligible to compete. Thus, the less inclusive the small business set-aside category, the fewer contractors are eligible to compete for the contract. We, therefore, hypothesize that restriction on competition is expected to increase the price per ton and price per container of waste removal.

H5a: Less inclusive small business set-asides (i.e., 8(a) Sole Source, 8(a) Competed, HUBZone, and SDVOSB) result in a higher price per ton than the more inclusive small business set-aside (i.e., Total Small Business).

H5b: Less inclusive small business set-asides (i.e., 8(a) Sole Source, 8(a) Competed, HUBZone, and SDVOSB) result in a higher price per container than the more inclusive small business set-aside (i.e., Total Small Business).

Among the less inclusive small business set-aside categories, one category is particularly exclusive: 8(a) Sole Source. In this situation, the contracting officer can choose not to compete the requirement at all; instead he or she can simply award the contract to an eligible 8(a) contractor. Therefore, we hypothesize that without competition, the price per ton and price per container of waste removal is expected to increase.

H6a: Among the less inclusive small business set-asides, the sole source set-aside (i.e., 8(a) Sole Source) results in a higher price per ton than the competed set-asides (i.e., 8(a) Competed, HUBZone, and SDVOSB).

H6b: Among the less inclusive small business set-asides, the sole source set-aside (i.e., 8(a) Sole Source) results in a higher price per container than the competed set-asides (i.e., 8(a) Competed, HUBZone, and SDVOSB).

Performance-Related Hypotheses

Similar to our first hypothesis, in our seventh hypothesis we seek to determine if each of the ISWM- and contracting-related variables affect contractor performance. Because the ISWM-related variables (i.e., tons of waste, number of containers, and wage rate) were provided to the contractors early in the acquisition process, were understood prior to vendor bidding, and tend to remain stable throughout the life of the contract, we do not expect to find that ISWM-related variables significantly affect performance.

H7a: ISWM-related variables do not affect contractor performance.

On the other hand, because adequate competition is known to simultaneously decrease price and increase performance, we do expect to find a significant relationship between the contracting-related factors (i.e., small business set-asides and number of offers) and performance.

H7b: Contracting-related variables affect contractor performance.

Methods

To test the hypotheses, we use three different statistical methods. We first describe the price-related methods and then move to the performance-related methods.

Sequential Multiple Regression

For H1 to H3, we use sequential multiple regression to determine the amount of variance in price (i.e., increase in R2) captured by each variable.

> In sequential regression (sometimes called hierarchical regression), independent variables enter the equation in an order specified by the researcher. Each IV (or set of IVs) is assessed in terms of what it adds to the equation at its own point of entry. … The researcher normally assigns order of entry of variables according to logical or theoretical considerations. (Tabachnick & Fidell, 2007, p. 138)

To test the amount of variance in price each IV captures, we entered them in sequence with the hypotheses. Specifically, we made five groups of predictors: Group

$k = 1$ consists of v_1 = Number of Tons; Group $k = 2$ consists of v_2 = Number of Containers; Group $k = 3$ consists of v_3 = Wage Rate; Group $k = 4$ consists of $v_4, \ldots, v_8$ = Small Business Set-Aside categories; and, Group $k = 5$ consists of v_9 = Number of Offers.

Accordingly, we perform $k = 1 \ldots 5$ linear regressions given by Equation 1:

$$p_k = a_k + \sum_{i \in \text{Group } 1\ldots k} b_{ik} v_i + e_k, \quad \forall k = 1 \ldots 5 \tag{1}$$

where, at the k-th stage in the sequence: p_k is dependent variable Price; a_k is the intercept regression coefficient; b_{ik} is the slope regression coefficient associated with dependent variable i; v_i is the value of the i-th variable; and, e_k is the error term.

Note that in this sequential approach, the group order in which the new variable(s) are added to explain the DV matters. Given our knowledge of the problem, we posit that Number of Tons should have the leading role, and so on. We later revise this assumption based on the results.

Also, like any regression analysis, certain assumptions about the data were met prior to performing the regressions. First, the Small Business Set-Aside categories are dummy variables. We exclude Full & Open Competition in order to compare the set-asides to full competition. Also, we started with 63 observations; however, in the course of testing our assumptions, we removed 6 outliers, thus reducing our useful observations to n = 57. Normality, linearity, and homoscedasticity of the residuals were verified. Multicollinearity was ruled out and the errors were deemed to be independent (i.e., non-correlated).

Wilcoxon Rank Sum Test

For H4 through H6, we use the Wilcoxon Rank Sum Test to determine if the median prices of the groups are statistically different. The Wilcoxon Rank Sum Test is the non-parametric equivalent of the independent t-test, which is used to determine if there is a statistically significant difference between the means of two unrelated groups. We use this non-parametric test because the price for each of the categories was not normally distributed; however, the general shape of the distributions for each group were the same. The null hypothesis for this test is that there are no differences in price between the groups being compared—that they have equal medians. The groups we compare and associated results are displayed in the next section.

Ordered Logistic Regression

For H7a and H7b, we use ordered logistic regression to determine whether or not the ISWM- and contracting-related variables affect contractor performance. Ordered logistic regression is appropriate given the categorical (i.e., non-continuous) nature of the DVs. The categorical nature of the performance scale makes it inappropriate for multiple regressions. Ordered logistical regression is like the more typical binary logistic regression in that it makes probabilistic predictions that an observation belongs in a given category; however, ordered logistic regression is appropriate for outcomes with multiple (vice the binary two) categories. Ordered logistic regression uses a series of equations to determine the probability that the observation is above the first category (i.e., above Unsatisfactory), above the second category (i.e., above Marginal), and so on. Equation 2 shows this multiple-category approach. The right-hand side of the equation represents the more common logistic regression (here *u* represents a linear regression calculation involving any number of predictors). The equation predicts the probability that the actual outcome *Y* exceeds category *j*.

$$\Pr\{Y > j\} = \frac{1}{1 + e^{-u}}, \quad \forall j \qquad (2)$$

With the hypotheses specified and the methods described, we turn to the results and implications.

Results

Price-Related Results

Sequential Multiple Regression Results

The results of the sequential multiple regressions are provided in Table 4. When using price as the DV, we found that ISWM variables account for 45% of the variance in price, while contracting-related variables accounted for 32%.[1] Further, the total η^2 for the ISWM-related variables was 0.21, while the total η^2 for the contracting-related variables was 0.14. These results suggest that the ISWM service-related variables (tons of waste, number of containers, and wage rate) influence price more than the contracting-related variables (small business set-asides and number of offers). Thus, H1 is supported. This is welcome news for buying organizations, as most desire to make their processes as efficient as possible in order to have minimal effect (if any) on price.

Testing the relative effects the service-related variables have on price, we find that number of tons (η^2 = 0.02) does not have a greater effect on price than number of containers (η^2 = 0.18). Thus, H2a is not supported. However, H2b is supported, as number of containers (η^2 = 0.18) has a greater effect on price than wage rate (□2 = 0.01). These results suggest that the largest service-related price driver is the number of containers, followed by the number of tons of waste, and finally wage rate. Clearly, organizations receiving the ISWM service should examine the number of containers they are using, as reducing containers may significantly reduce price.

Testing the relative effects the contracting-related variables have on price, we find that the small business set-asides (total η^2 = 0.11) have a greater effect on price than the number of offers received (η^2 = 0.03). Thus, H3 is supported. This result is intuitive, but also important in the sense that buying organizations cannot simply reduce price by stirring up competition. Buying organizations should understand the price premium they can expect to pay for meeting certain socio-economic goals so they can make informed acquisition decisions.

---

[1] To account for the fact that there is declining available DV variance the later a variable is input into the regression, we performed a second sequential multiple regression whereby the contracting-related variables were entered first. In this analysis, we found that the ISWM-related variables accounted for 40% of the variance in price, and the contracting-related variables accounted for 37% of the variance.

Table 4. Sequential Regression Results

Sequence (*k*)	Variables (*i*) in Sequential Regression	Regression Coefficient	Standard Coefficient	η^2	Total R^2	Change in R^2
1	Number of Tons	59.62***	.60***	.36	.36	--
2	Number of Tons	36.25**	.34**	.09	.45	.09
	Number of Containers	802.19**	.44**	.15		
3	Number of Tons	35.83*	.34*	.09	.45	.00
	Number of Containers	810.13**	.44**	.15		
	Wage Rate	-2501.98 ns	-.04 ns	.00		
4	Number of Tons	36.31*	.34*	.08	.50	.05
	Number of Containers	654.84*	.36*	.08		
	Wage Rate	2839.61 ns	.04 ns	.00		
	8(a) Sole Source	39480.93 ns	.08 ns	.00		
	8(a) Competed	124260.60 ns	.16 ns	.02		
	HUBZone	158374 ns	.20 ns	.03		
	SDVOSB	-84663.86 ns	-.06 ns	.00		
	Total SB	35033.08 ns	.08 ns	.00		
5	Number of Tons	13.75 ns	.17 ns	.02	.77	.27
	Number of Containers	776.41***	.54***	.18		
	Wage Rate	-7178.45 ns	-.13 ns	.01		
	8(a) Sole Source	121095.20**	.33**	.06		
	8(a) Competed	97168.88 ns	.16 ns	.01		
	HUBZone	122387.10 ns	.20 ns	.03		
	SDVOSB	-61995.22 ns	-.06 ns	.00		
	Total SB	43398.03 ns	.13 ns	.01		
	Number of Offers	18275.81*	.30*	.03		

*p < .05; **p < .01; ***p < .001; ns stands for non-significant
R^2 is the % of variance in the DV (price) that is explained by the IVs.
η^2 is the % of variance in the DV (price) that is explained by each IV.

Wilcoxon Rank-Sum Test Results

The results of the Wilcoxon Rank-Sum Test did not support H4, H5, or H6. Table 5 illustrates the results.

H4a and H4b test whether there is a difference in price per ton and price per container, respectively, between small business set-asides and full and open competition contract awards. For H4a, contracts solicited as small business set-asides did not result in significantly higher prices per ton than contracts that were solicited using full and open competition. In these data, the price per ton for small business set-asides is $198/ton, while the price per ton for all others is $132/ton. Although the difference may appear large, it is not statistically significant.

However, there were significant differences in the price per container between small business set-asides and full and open competitions. These results support H4b. The price per container for small business set-asides is $2,101/container, while the price per container for full and open competitions is $1,407/container. In these data, the buying organization appears to be paying approximately $700 more per container on small business set-asides. This result again calls for the organizations receiving the service to carefully examine the number of containers they are using, particularly given the fact that the differences in price per ton were not significant. In other words, it is not the amount of waste disposed that affects the price difference between small business set-asides and full and open competition; rather it is the number of containers being serviced.

H5a and H5b test the notion that less inclusive small business set-asides would result in higher price per ton and price per container, respectively, than a simple Total Small Business set-aside. In these data, the price per ton for less inclusive small business set-asides is $156/ton, while the price per ton for Total Small Business set-asides is $249/ton. Although not statistically significantly different, the results are actually counterintuitive, with the more inclusive category having a higher price per ton of waste removal. Thus, H5a is not supported.

For H5b, the less inclusive set-asides did result in a higher price per container ($2,132/container) than the more inclusive set-asides ($2,063/container), however the difference was not statistically significant. Thus, H5b is not supported.

Finally, H6a and H6b test the notion that sole source small business set-asides would result in higher price per ton and price per container, respectively, than the other less inclusive small business set-asides. In these data, the price per ton for 8(a) sole source set-asides is $139/ton, while the price per ton for the other less inclusive small business set-asides is $174/ton. Although not statistically significant, these results are also counterintuitive—the sole source price per ton is less than the competed price per ton amongst less inclusive small business set-aside categories. Thus, H6a is not supported.

The results of H6b were the same as H6a. Again, although not statistically significantly different, the sole source price per container ($2,116/container) is slightly less than the competed price per ton ($2,156/container) amongst less inclusive small business set-aside categories. Thus, H6b is not supported.

The results of H5 and H6 suggest that once the buying organization has chosen to solicit the requirement using a small business set-aside, the type of set-aside does not affect price per ton or price per container. This information is critical to the buying organization, as they often try to spread their budgets among the different set-aside categories in order to meet Small Business Administration goals. Using less inclusive set-asides may help organizations meet their SBA goals faster, assuming the organization is able to meet the requirements for fair and reasonable pricing (see FAR 19.502-2(b)).

Table 5. Wilcoxon Rank-Sum Test Results*

Hypothesis	Group 1	Group 2	Result
	Contracts with Small Business Set-Aside Categories 8(a) Sole Source 8(a) Competed HUBZone SDVOSB Total Small Business Set-Aside	Contracts with No Small Business Set-Aside Full & Open Competition	
H4a Price/Ton	$198.39 *n*=35	$131.61 *n*=14	ns
H4b Price/Container	$2,100.78 *n*=31	$1,406.78 *n*=13	p<.01
	Less Inclusive Small Business Set-Aside Categories 8(a) Sole Source 8(a) Competed HUBZone SDVOSB	More Inclusive Small Business Set-Aside Category Total Small Business Set-Aside	
H5a Price/Ton	$155.76 *n*=19	$249.02 *n*=16	ns
H5b Price/Container	$2,132.03 *n*=17	$2,062.82 *n*=14	ns
	Sole-Source Small Business Set-Aside Category 8(a) Sole Source	Competed Small Business Set-Aside Categories 8(a) Competed HUBZone SDVOSB	
H6a Price/Ton	$139.07 *n*=10	$174.30 *n*=9	ns
H6b Price/Container	$2,115.51 *n*=10	$2,155.63 *n*=7	ns

* Note. *n* refers to the sample size used in each case. The result column indicates the significance level or non-significant (ns).

Performance-Related Results

Interestingly, we found that neither the ISWM- nor contracting-related variables affected contractor performance. These results support H7a, but not H7b. The results suggest that there are no differences in quality, cost, schedule, or management performance[2] based on (1) the amount of the service required (i.e., large versus small tonnage, large versus small number of containers), (2) the prevailing wage rate in a given area, (3) whether the requirement was solicited and awarded using a small business set-aside or full and open competition, or (4) the size of the competition (i.e., large number of offers versus small number of offers). It should be noted that only 32 of the remaining 57 observations had contractor performance data. More of these data would be required to confirm these results.

---

[2] Given the few performance ratings available for small business subcontracting, the ordered logit could not converge, thus small business subcontracting was removed from the individual DV analysis. We also tested the combined average performance score across all five performance categories (to include small business subcontracting); however, the results were not different from those reported. A larger set of performance data is needed to confirm these results.

A summarized version of our hypotheses and their related results are presented in Table 6.

Table 6. Summary of Hypotheses and Results

Hypothesis	Description	Supported?
H1	ISWM service-related variables have a greater effect on price than contracting-related variables.	Y
H2a	Tonnage of waste has a greater effect on price than number of containers.	N
H2b	Number of containers has a greater effect on price than wage rate.	Y
H3	Small business set-asides have a greater effect on price than number of offers.	Y
H4a	Small business set-asides result in a higher price per ton than full and open competition.	N
H4b	Small business set-asides result in a higher price per container than full and open competition.	Y
H5a	Less inclusive small business set-asides (i.e., 8(a) Sole Source, 8(a) Competed, HUBZone, and SDVOSB) result in a higher price per ton than the more inclusive small business set-aside (i.e., Total Small Business).	N
H5b	Less inclusive small business set-asides (i.e., 8(a) Sole Source, 8(a) Competed, HUBZone, and SDVOSB) result in a higher price per container than the more inclusive small business set-aside (i.e., Total Small Business).	N
H6a	Among the less inclusive small business set-asides, the sole source set-aside (i.e., 8(a) Sole Source) results in a higher price per ton than the competed set-asides (i.e., 8(a) Competed, HUBZone, and SDVOSB).	N
H6b	Among the less inclusive small business set-asides, the sole source set-aside (i.e., 8(a) Sole Source) results in a higher price per container than the competed set-asides (i.e., 8(a) Competed, HUBZone, and SDVOSB).	N
H7a	ISWM-related variables do not affect contractor performance.	Y
H7b	Contracting-related variables affect contractor performance.	N

Conclusion

This first-of-kind study empirically tested the impact small business set-asides have on contract price and contractor performance. When all ISWM service- and contracting-related variables are included in a regression, we find that the number of containers (a service-related variable) has the largest effect on price. This result is particularly important, as it suggests that the USAF may be able to significantly reduce the price of their ISWM contracts simply by managing the number of containers that must be serviced on each base.

Two contracting-related variables, 8(a) Sole Source set-aside and number of offers, also significantly affect price. Ironically, the results suggest that as the number of offers increases, the total price also increases. These results are particularly counterintuitive, as the ISWM requirement would typically be subject to the lowest cost technically acceptable source selection method, where price is the main determinant of award.

Interestingly, we find no differences in price per ton between (1) small business set-asides and full and open competition, (2) less inclusive small business set-asides and the more inclusive Total Small Business set-aside category, and (3) the 8(a) Sole Source set-aside category and the less inclusive competed set-asides. Using the same comparison categories, we find only one difference in price per container: between small business set-asides and full and open competition. These results once again highlight the importance of number of containers as a price driver, and suggest that buying organizations can choose to target their small business set-asides without significantly affecting price per ton or price per container.

It should be noted that while most differences in prices are not statistically significant, there are still differences. Knowing the median prices paid across USAF bases, as well as the difference in the median prices between comparison categories, may help acquisition teams craft their strategies and understand whether or not received proposals represent a relatively good or a relatively bad deal, as compared to historical prices paid.

Like all research, there were limitations to our analyses. Data limitations do not allow us to account for other factors that may affect the price and performance of the ISWM service, such as distance from the Air Force base to the landfill, the cost to dispose of waste in a given geographical area, and the size and capacity of the trucks being used to pick up and dispose of the waste. We suggest those variables be captured for future analyses.

Data limitations also limit the generalizability of the Wilcoxon Rank-Sum Test results. For adequate statistical power, each comparison group should contain at least 15 observations. That criterion was only met for five of the 12 groups. Finally, more contractor performance data are needed to reach more accurate conclusions concerning contractor performance and ISWM- and contracting-related variables.

References

Achillas, C., Moussiopoulos, N., Karagiannidis, A., Banias, G., & Perkoulidis, G. (2013). The use of multi-criteria decision analysis to tackle waste management problems: A literature review. *Waste Management & Research, 31*(2), 115–129.

Apte, A., Rendon, R. G., & Salmerón, J. (2011). An optimization approach to strategic sourcing: A case study of the United States Air Force. *Journal of Purchasing & Supply Management, 17*(4), 222–230.

Arribas, C. A., Blazquez, C. A., & Lamas, A. (2010). Urban solid waste collection system using mathematical modelling and tools of geographic information systems. *Waste Management & Research, 28*(4), 355–363.

Boskovic, G., Jovicic, N., Jovanovic, S., & Simovic, V. (2016). Calculating the costs of waste collection: A methodological proposal. *Waste Management & Research, 34*(8), 775–783.

GAO. (2016, October). *Federal procurement: Smarter buying initiatives can achieve additional savings, but improved oversight and accountability needed* (GAO-17-164). Washington, DC: Author.

Henderson, B. D. (1975). The coming revolution in purchasing. *Journal of Purchasing and Materials Management, 11*(2), 44–46.

Kraljic, P. (1983). Purchasing must become supply management. *Harvard Business Review, 61*(5), 109–117.

Monczka, R. M., Handfield, R. B., Giunipero, L. C., & Patterson, J. L. (2015). *Purchasing and supply chain management.* Independence, KY: Cengage Learning.

Office of Management and Budget (OMB). (2012, December 5). *Improving acquisition through strategic sourcing* [Memorandum]. Retrieved from https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-02_0.pdf

Office of Management and Budget (OMB). (2014, December 4). *Transforming the marketplace: Simplifying federal procurement to improve performance, drive innovation, and increase savings*. Retrieved from https://www.whitehouse.gov/blog/2014/12/04/transforming-marketplace-simplifying-federal-procurement-improve-performance-drive-i

Rendon, R. G. (2005). Commodity sourcing strategies: Processes, best practices, and defense initiatives. *Journal of Contract Management, 3*(1), 7–20.

Rendon, R. G., & Templin, C. R. (1992, July). Corporate procurement strategy: An analysis of supply line management. *Contract Management, 32*(7), 18–25.

Sharkey, J. (2015, November 4). *Buying as one through category management* [Presentation slides]. Retrieved from http://mvs013-020.directrouter.com/~sameorg/images/stories/continuing_education/sbc_2015/1545-1645_room217.pdf

Solano, E., Ranjithan, S. R., Barlaz, M. A., & Brill, E. D. (2002). Life-cycle-based solid waste management. I: Model development. *Journal of Environmental Engineering, 128*(10), 981–992.

Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA*. Belmont, CA: Thomson/Brooks/Cole.

USA Spending. (2016). Overview of awards—FY 2016. Retrieved from https://www.usaspending.gov/Pages/Default.aspx

THIS PAGE INTENTIONALLY LEFT BLANK

Panel 17. Software Trends and Issues

Thursday, April 27, 2017	
1:45 p.m. – 3:15 p.m.	**Chair: Captain Kurt Rothenhaus, USN,** PMW-160 Program Manager, Navy Tactical Networks
	Software Productivity Trends and Issues
	David M. Tate, Institute for Defense Analyses
	Transformation of Test and Evaluation: The Natural Consequences of Model-Based Engineering and Modular Open Systems Architecture
	Nickolas H. Guertin, Naval Surface Warfare Center, Division Carderock CAPT Gordon Hunt, USN, Navy Reserve
	Decision-Based Metrics for Test and Evaluation Experiments
	Dashi Singham, Naval Postgraduate School

Captain Kurt Rothenhaus, USN—assumed command of Space and Naval Warfare Systems Center Pacific (SSC Pacific) on December 17, 2013. SSC Pacific consists of more than 4,800 civilian and military personnel, charged with the research, development, fielding, and support of advanced technologies for command, control, communications, computers, intelligence, surveillance, and reconnaissance (C4ISR), cyber, and space capabilities.

A native of New York City, CAPT Kurt Rothenhaus received his commission upon graduating from the University of South Carolina. He holds a Master of Science degree in Computer Science and a PhD in Software Engineering from the Naval Postgraduate School and transferred into the Engineering Duty Officer community in 2003.

CAPT Rothenhaus' operational assignments include USS *Fife* (DD 991), USS *O'Brien* (DD 975), Destroyer Squadron Fifteen, and Combat Systems/C5I Officer on USS *Harry S. Truman* (CVN 75). Additionally, he served in Baghdad, Iraq, developing counter-insurgency and reconstruction systems for the Army Corps of Engineers in 2006.

His acquisition assignments include project manager, SSC Pacific, and various acquisition leadership roles in Program Executive Office, Command, Control, Communication, Computers, and Intelligence (PEO C4I) to include Future Command and Control Systems assistant program manager in the Navy Command and Control Program Office (PMW-150), and assistant program manager for the Consolidated Afloat Network Enterprise Services (CANES) in the Navy Tactical Networks Program Office (PMW-160). Additionally, he served as the deputy program manager for Navy Communications and GPS Program Office (PMW/A-170) from September 2011 to October 2013.

CAPT Rothenhaus has been recognized with the A. Bryan Lasswell, National Defense Industrial Association Award in 2007 for technology innovation, and was a 2008 Navy and Marine Corps Leadership Award winner while serving aboard USS *Harry S. Truman* (CVN 75). His personal awards include the Meritorious Service Medal, Joint and Navy and Marine Corps Achievement Medal, Navy and Marine Corps Commendation medals, and various service and campaign awards.

Software Productivity Trends and Issues

David M. Tate—joined the research staff of the Institute for Defense Analyses' Cost Analysis and Research Division in 2000. Since then, he has worked on a wide variety of resource analysis and quantitative modeling projects related to national security. These include an independent cost estimate of Future Combat Systems development costs, investigation of apparent inequities in Veterans' Disability Benefit adjudications, and modeling and optimization of resource-constrained acquisition portfolios. Dr. Tate holds bachelor's degrees in philosophy and mathematical sciences from Johns Hopkins University, and MS and PhD degrees in operations research from Cornell University. [dtate@ida.org]

Abstract

The Department of Defense is experiencing an explosive increase in its demand for software-implemented features in weapon systems. The combination of exponential increases in computing power and similar advances in memory density and speed has made software-mediated implementation of system features increasingly attractive. In the meantime, defense software productivity and industrial base capacity have not been growing as quickly. Do we have an impending bottleneck? If so, what are the management implications?

Malthus on Software

The Scottish cleric and economist Thomas Robert Malthus famously noted that, when there is enough food to go around, population growth is exponential. Since Malthus could not envision any means whereby food production could also grow exponentially, given the constraints of arable land and property ownership, he predicted that the inevitable result would be a population limited by recurring poverty and starvation.

Malthus was wrong about food, at least so far, but could he be right about national security software? Any time you have an exponential growth in demand without a commensurate exponential growth in supply, demand will soon be frustrated. Rapidly growing demand for new software, combined with the need to maintain the code going forward, places considerable stress on the productive capacity of the defense software industrial base. The ability to keep up will depend on just how fast demand is growing, how quickly the Department of Defense (DoD) can grow the industrial base, and how quickly the productivity of individual software developers improves over time. To know whether we should worry, we need to look at each of those factors.

How Fast Is Defense Demand for Software Growing?

It is surprisingly difficult to find historical and current data on the demand for software in defense systems. However, there are some strong indicators available:

- The National Research Council (2010) wrote that "the extent of the DoD code in service has been increasing by more than an order of magnitude every decade, and a similar growth pattern has been exhibited within individual, long-lived military systems." One order of magnitude per decade is approximately 25% annual growth.
- The Aerospace Vehicle Systems Institute (2017) states that source lines of code (SLOC) in aircraft (both military and commercial) has been doubling approximately every four years. This corresponds to an annual growth rate of ~18%.
- The Army (2011) estimated that the volume of code under Army depot maintenance (either post-deployment or post-production support) had

increased from 5 million to 240 million SLOC between 1980 and 2009. This corresponds to ~15% annual growth.

- Dvorak (2009) stated that National Aeronautics and Space Administration unmanned space systems SLOC have also increased by an order of magnitude every 10 years, with manned systems SLOC growing even faster.

Taken together, these suggest an annual growth rate of at least 15% for the amount of software being developed and maintained for defense purposes, with 25% or more annual growth possible. Annual growth of 15–25% means doubling every three to five years, on top of which is the added workload of maintaining the growing base of deployed code.

In order to forecast future demands for new code and software maintenance, we also need to know the current size of the code base and the current annual demand. The most recent demand estimate we were able to find (Chao, 2006) concluded that the 2006 requirement for national security software was about 35 million lines of new code and 25 million lines of maintenance code. We can apply the “20% per year” rule of thumb for maintenance effort to infer a deployed 2006 base of about 125 million lines of code. We will base our analysis on those assumptions: 125 million source lines of code (MSLOC) under maintenance in 2006, 35 MSLOC of new code required in 2006, and annual demand for new code growing at 15% annually from that time forward. For maintenance effort, we assume that annual maintenance effort on the installed base is equivalent to 20% of the development effort of the base, and that half of the maintenance effort results in more new code to be maintained.[1] In addition, some fraction of the installed code base is retired every year. We will assume that 10% of the installed base is retired each year, exactly offsetting the new code generated by maintenance. As we will see, the conclusions of this investigation are not sensitive to the exact parameter values chosen here.

Figure 1 shows the projected growth in annual demand for defense software under these assumptions, separated into new code and maintenance of existing code. Bear in mind that this is a projection of *unconstrained* demand—how much the DoD is expected to want to buy, if it is available at prices comparable to historical prices.

[1] Jones (2013) estimates the maintenance costs of a nominal 1000-function point application at closer to 40% per year over the first five years. Using that estimate would result in a smaller 2006 deployed code base estimate, but much faster growth in that base in subsequent years.

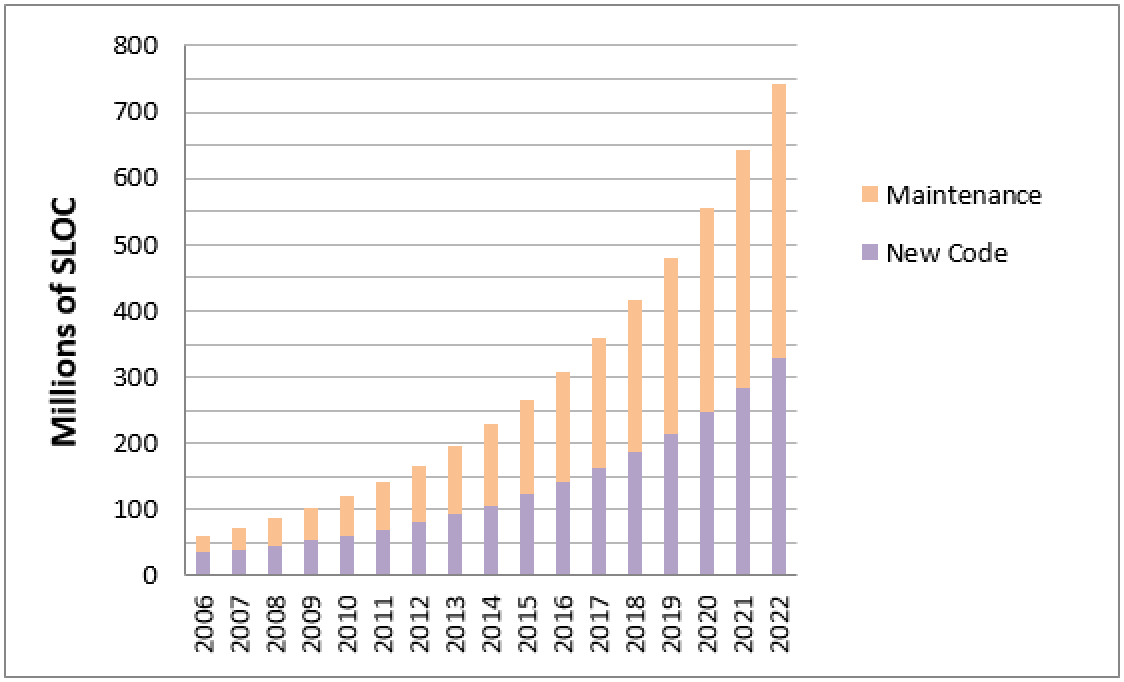


Figure 1. Forecast of DoD Software Demand

It is worth noting that, under these assumptions, the total size of the deployed code base under maintenance is projected to be more than 1 billion SLOC by 2018, and more than 3 billion SLOC by 2025. Figure 1 shows only the new effort each year, not the deployed base.

The Supply of Defense Software

Chao (2006) estimated both the size of the defense software workforce and the productivity of that workforce. The productive capacity of the industrial base is the product of those two factors. We will attempt to estimate each in turn using available data. For purposes of this analysis, we will accept Chao's estimates that there were 68,000 cleared software developers in 2006, capable of producing 75 MSLOC per year. That implies a productivity at that time of roughly 1100 SLOC per developer in 2006, or (equivalently) 900 developers required per MSLOC, as our baseline.

The Size of the Workforce

How quickly might the defense software workforce be growing? The Bureau of Labor Statistics estimates that from 2010 to 2015, total employment of software developers[2] grew almost 30%, or about 5.3% annually. However, they forecast that rate to decline sharply going forward, averaging only about 1.6% per year over the decade of 2014–2024 (BLS, 2017). The defense software industrial base will need to grow more quickly than that to keep pace with established demand growth.

[2] BLS occupation codes 15-1132 (software developers, applications) and 15-1133 (software developers, system software), total employment as of May 2010 and May 2015 (BLS, 2017).

Any scarcity of cleared software talent should translate into rising salaries and benefits for workers with those skills, providing incentive for more and more workers to enter the industry. In a free and liquid market, we would expect this to happen fairly quickly. Unfortunately, some aspects of this particular market might be problematic. The first is the requirement that workers be U.S. citizens with security clearances. This not only dramatically restricts the pool of potential entrants, but it also creates a licensure bottleneck for individuals seeking to join the labor force. There is currently a backlog of half a million unfinished security background checks, and the time required to get through the process is increasing steadily. In addition, over the past few years the total size of the cleared workforce has contracted by ~25% in reaction to high-profile spills of classified information. Defense software employers are also facing tough competition from the private sector, which is experiencing an explosion of demand for software to power the expanding role of the Internet in daily life. Of course, other industries can supplement U.S. graduates with offshore or immigrant labor—a solution unavailable to the defense sector under current regulations.

Another barrier to market corrections is that the most urgent scarcities seem to be at the high end of the experience scale. Chao (2006) found that (at least in 2006) there was no general shortage of programmers, but there was already a significant shortage (with corresponding salary premium) of relatively senior software project managers, architects, and developers. At the tip of the pyramid, they cited a cadre of 500–600 "elite" individuals who play a disproportionate role in project success. If rising compensation for senior talent begins to cause an increased growth rate in software degrees, we will not see that begin to alleviate the crunch in senior talent and elite individuals for at least another 10 years.

Finally, it is not clear that the DoD *wants* the market to correct itself through increases in compensation. Contractor labor rates are closely monitored by the DoD, and the government pushes back when they rise too quickly. Senior software talent in the general economy can be as highly compensated as senior management executives. Arrington (2010a) reported that "[a Google employee] was recently offered a counter offer he couldn't refuse (except he did). He was offered a 15% raise on his $150,000 mid level developer salary, quadruple the stock benefits and … wait for it … a $500,000 cash bonus to stay for a year. He took the Facebook offer anyway." (Note that $150,000 for a mid-level developer is already well above industry norms.) Arrington (2010b) also reported that Google had paid a top software engineer $3.5 million to turn down an offer from Facebook. Allowable defense contractor labor costs are capped; companies choosing to pay salaries over those caps must take the difference out of profit. This provides a strong disincentive to paying market rates for top talent within the defense world.

On the supply side, what does the educational pipeline for software look like? The number of bachelor's degrees conferred each year in computer and information sciences has shown a striking cyclical pattern over the past four decades (see Figure 2). The general trend has been a baseline increase of ~1000 degrees per year, with superimposed boom and bust cycles. We are currently on the upswing of a boom cycle, with more than 60,000 degrees conferred per year.

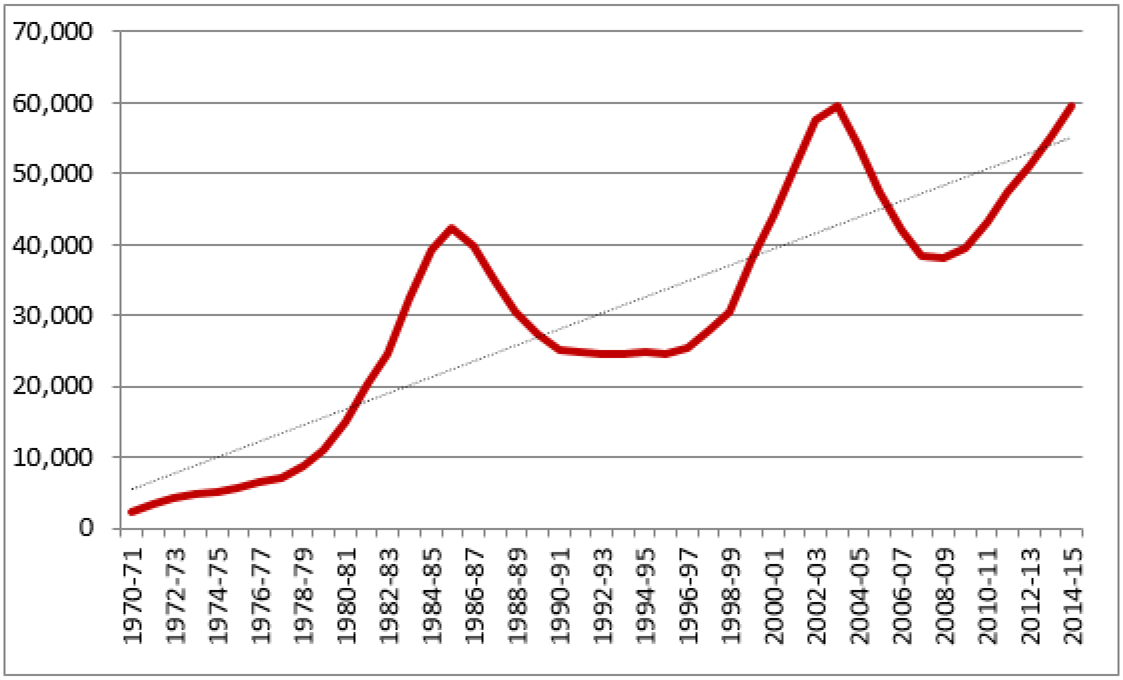


Figure 2. Annual Computer Science and Information Sciences Bachelor's Degrees Conferred

In addition to this pool of potential defense software developers, the educational pipeline for software developers also includes nontraditional educational options. More than 16,000 students graduated from "coding boot camp" programs in 2015, and that number has been growing rapidly over the few years that such programs have existed (Lauerman, 2015).

This suggests that there are as many as 80,000 potential developers graduating per year. In 2006, the cleared workforce made up 7% of the overall software workforce. Again being optimistic, if 10% of new graduates (college and boot camp combined) end up in the cleared software workforce, that would currently be about 8,000 per year, which could grow to 10,000 per year in a couple of years. This corresponds to between 5% and 10% annual growth. For purposes of our baseline analysis, we will assume annual workforce growth of 5%, comparable to recent growth in software developers and well above the forecast national average for the software industry.

As noted above, in 2006, there were roughly 68,000 cleared software developers in the defense industrial base. If we assume 5% annual growth in the national security software developer workforce starting in 2006, that would translate to about 120,000 people today, reaching 150,000 by 2023. Figure 3 shows this projected growth over time.

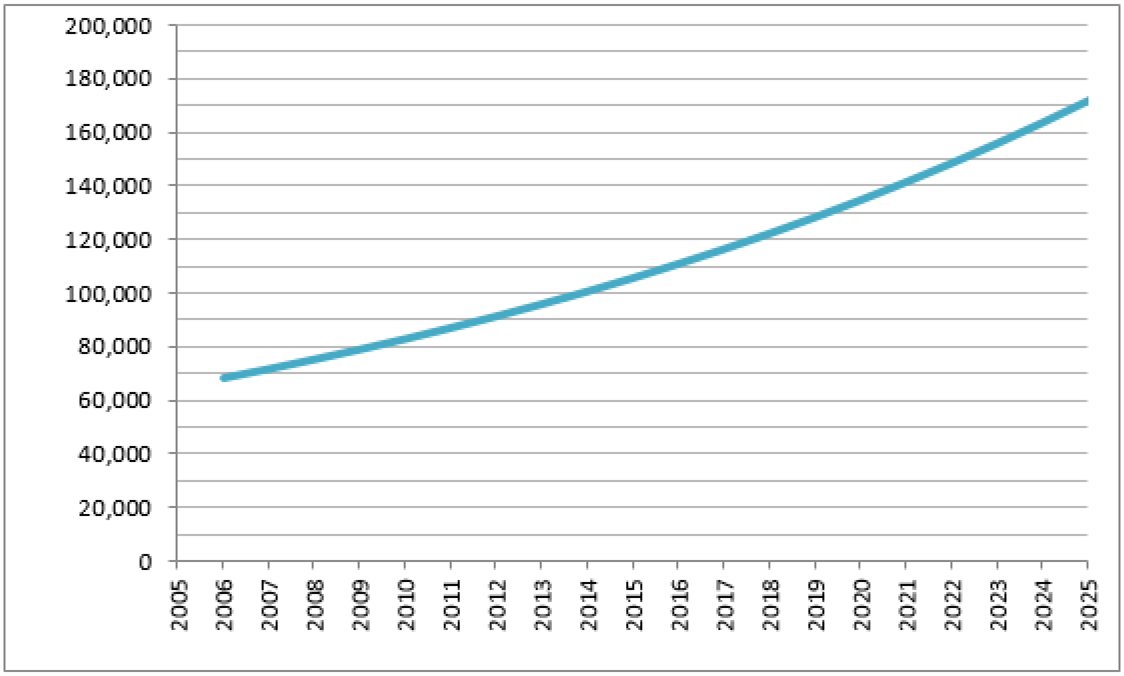


Figure 3. Forecast Cleared Software Workforce Size

The Productivity of Defense Software Developers

Malthus was wrong about hunger in England in large part because the technology for food production improved enormously over the next few centuries, making individual farmers much more productive and bringing marginal land into productive use. Could defense software development see (or already be seeing) a similar explosion in individual productivity that would be enough to make up for the slower growth of the labor force?

In 2000, Jones estimated defense software productivity at 4.2 function points (FP) per staff month (SM); in 2013, his estimate was 6.75 FP/SM. That corresponds to just under 4% annual productivity improvement. This is in line with other historical estimates of software productivity growth. For example, Longstreet (2001) estimated ~4% annual productivity growth (FP per hour) from 1970 to 2000 industry-wide. These estimates are based on FP, rather than on MSLOC. Since the number of FP per line of code has been growing historically (Jones, 2013), productivity growth in terms of MSLOC would be somewhat lower, but we will optimistically estimate MSLOC productivity growth at 4% as well.

Of course, the DoD may not yet have realized all of the productivity enhancement that can be had using current technology. We discuss these at greater length in the Recommendations section.

Supply vs. Demand

We now have all of the pieces we need for a back-of-the-envelope comparison of forecast productive capacity versus unconstrained demand. Figure 4 shows that, even given the generally optimistic assumptions we have made, we have already passed the point of being able to produce and maintain all of the software that the DoD would like. According to this forecast, the DoD will soon also reach the point of neither being able to produce all of the new code desired (without maintenance), nor to maintain all existing code (with no new development). The projected 2020 workforce of 135,000 developers would be less than half of the 290,000 developers required to write and maintain all of the code desired up to that point.

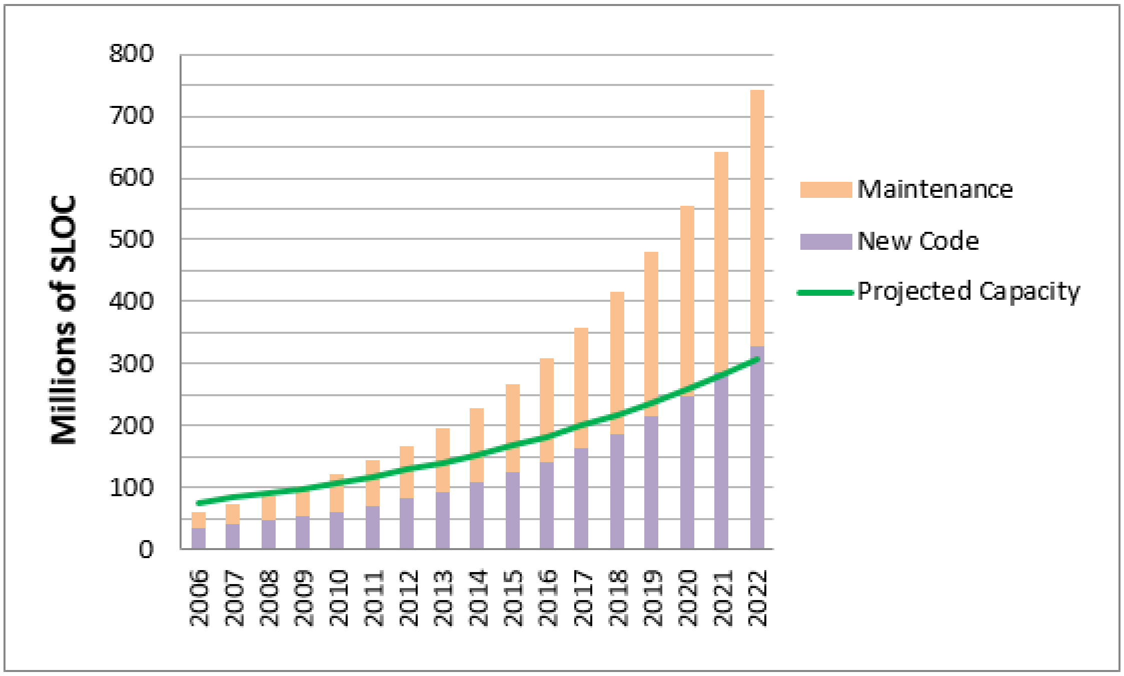


Figure 4. Forecast Supply vs. Unconstrained Demand

Revisiting the assumptions behind this forecast, we have assumed:

- 15% annual growth in demand for new code
- 5% annual defense software workforce expansion
- 4% annual productivity growth
- A workforce of 68,000 in 2006
- Demand for 35 MSLOC in 2006
- An installed base of 125 MSLOC in 2006
- Productive capacity of 75 MSLOC in 2006
- 20% annual maintenance effort
- 50% of maintenance resulting in new code
- 10% annual retirement of software in the base

Most of these assumptions could be fairly described as optimistic, based on historical data. Varying the parameters changes the details, but the shape of the situation remains the same. For example, if we assume that productivity growth post-2006 will be 8% instead of 4%, we get the picture in Figure 5. Software development is still capacity-constrained in this case, but not as severely. Conversely, if we keep productivity growth at 4% but allow the workforce to grow by 10% per year, we get the picture in Figure 6.

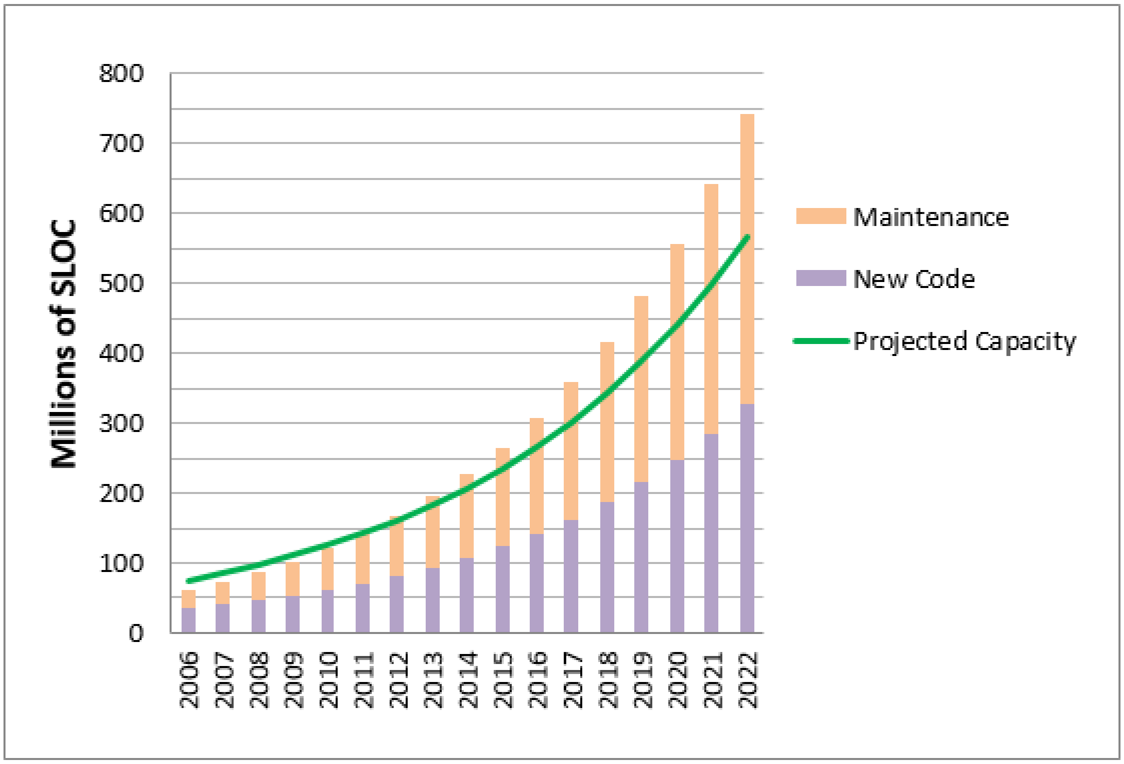


Figure 5. Supply vs. Demand at 8% Productivity Growth

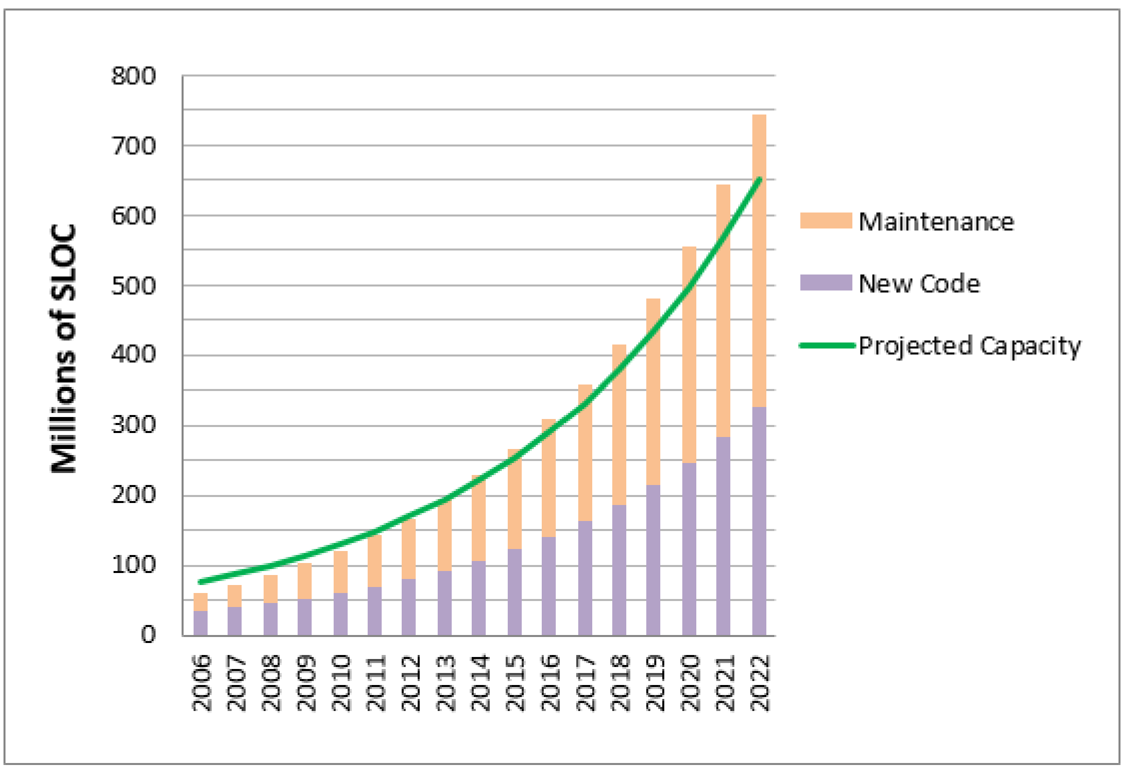


Figure 6. Supply vs. Demand Assuming 10% Workforce Growth

It goes without saying that the reverse is also true—if we assume 20% annual demand growth, or a 2006 installed base significantly larger than 125 MSLOC,[3] all of these pictures look much worse. Similarly assuming less optimistic values for the annual maintenance fraction (40%), or the proportion of maintenance that generates new code (>50%) (Galorath, n.d.), would lower the forecast capacity significantly.

If This Were Correct, Wouldn't We Have Noticed?

Is it really possible that we could be suffering a (possibly severe) shortage of software developers in the defense sector without anyone noticing? What symptoms should we look for?

Barnow, Trutko, and Piatak (2013) list 16 separate actions that employers might take that are indicative of a labor shortage. These include increased recruiting expenditures, increased use of overtime, new on-the-job training programs, relaxing minimum qualifications, and so forth. These are in addition to the operational symptoms of resource shortage, such as increased development times, lower-than-predicted staffing levels, and higher ratios of systems engineering/program management costs to touch labor costs.

Are these things happening in the defense sector? There is some evidence that they are.

- Chao (2006) found that senior software architects and project managers in the cleared software sector earned 50+% more than their counterparts in the general economy. They took this to indicate that there was already a shortage of those particular skills in the defense industrial base.
- Lucero (2009) found that many defense software positions were being filled by personnel with no formal software engineering training (on-the-job training).
- There are currently more than 10,000 job postings for software developers and software engineers at ClearanceJobs.com, which is more than half of all listings at that site (vacancies).
- Salaries for cleared information technology program/project managers rose 10% in one year between 2013 and 2014, faster than any other category and passing engineers as the highest-compensated cleared occupation group (salary rise) (ClearanceJobs.com, 2014).
- BLS estimates the national unemployment rate for technology professionals at only 2.9% (vacancies) (ClearanceJobs.com, 2014).
- Nearly half of recent ClearanceJobs.com survey respondents have been in their current job less than three years (churn) (Kyzer, 2017).

Barnow et al. (2013) also note that measuring occupational shortages is difficult, in part because occupational vacancy data are not generally available in the U.S. and available reporting uses job classification systems that are based on outdated industrial models and too coarse to be useful for many purposes. It would be very interesting to look

[3] Given that the Army alone claimed to have 240 MSLOC under sustainment in 2009, 125 MSLOC defense-wide in 2006 seems improbably low.

at (for instance) how the cost per staff month of defense software development has changed over the past decade, as reflected in the Software Requirements Data Reporting database.

What Are the Policy Options?

We identify several available short-term and long-term policy options associated with both the supply and demand for defense software.

Option 1: Moderate Demand

The obvious short-term solution to a scarcity of software productive capacity is to ask for less software. At the present time, it seems unlikely that the defense establishment would be willing or able to accomplish this. Software is viewed as vital to any hope of maintaining the United States' traditional technological advantage in military capability. A significant overall reduction in software demand would also require the several Services to cooperate effectively to optimize the allocation of software development capacity to the most important software-intensive programs. Given that the services struggle to allocate resources efficiently within and among their own acquisition portfolios, this seems like a stretch. The results, then, would be a less-efficient allocation of software resources to capabilities, an associated effective loss of software productivity, and failure to reap the potential benefits of software-mediated capabilities.

In the longer term, natural factors limit the growth in demand for software. Defense budgets do not grow without limit, so the exponential growth in software demand reflects, to some extent, substitution of software for other categories of expenditure—primarily analog hardware and human labor. There are natural limits to that process. Regardless of the underlying desire for software-mediated capabilities, the DoD cannot procure more software than the industrial base is able to provide.

Perhaps just as importantly, there is a tension between the size and complexity of the software in a system and how long it takes to develop that system. If rapid response to a rapidly changing world is one of the motivations for implementing capabilities in software, it makes no sense to pursue designs whose complex software will require 20+ years to design, build, and test. Prior analysis of the dependence of development cycle times on software content assumed development times unconstrained by industrial base issues (Tate, 2016). If Major Defense Acquisition Program/Major Automated Information System (MDAP/MAIS) software projects are now subject to chronic resource shortfalls, those past lead time estimates were optimistic. Increased demand for software-mediated functions thus has a twofold negative effect on schedules: first by adding work to the critical development path of each program, and second by starving the programs of the resources necessary to do the work on the critical path. From a policy perspective, it does not seem practical for the DoD or Congress to mandate reduced use of software overall, or limits on the amount of software in any one program. Not only would those policies be counterproductive, they would also be unenforceable and prone to wasteful gaming by the services and defense contractor base. Demand-side policy options would not seem to be helpful here.

Option 2: Grow the Workforce

From a policy perspective, there are several plausible mechanisms for increasing the growth rate of the defense software base:

- Encourage students to pursue software education, both through traditional college degrees and nontraditional (e.g., boot camp) training programs. Incentives could include low-interest loans, direct subsidies/scholarships, loan forgiveness, etc. These could be made contingent on a minimum tenure of employment in the defense sector.

- Invest in improving the throughput of the security clearance process, especially for software workers.
- Relax barriers to employing foreign nationals. The software industry has thoroughly globalized, but the defense sector is not permitted to take advantage of that at present. As we shall see below, there are ways of doing this implicitly that do not involved relaxing security standards.
- Allow contractors to pay true market salaries for software talent.

The first three of these options would tend to reduce the price of defense software by increasing supply, thus somewhat offsetting the investment required. Allowing higher salaries for key software professionals looks like it would tend to increase the cost of any given system—but it might not. It might improve efficiency and increase supply by enough to offset the higher cost per hour of that labor. It might also make it possible to have that system at all, or improve its quality, or permit the DoD to acquire it in time for it to be useful.

Option 3: Improve Productivity Dramatically

There have been multiple drivers of significant productivity improvement in the commercial software world over the past few decades. These include computer-aided software engineering (CASE) tools, automated test environments, improved programming languages,[4] agile (and similar) development processes, and modular open system architectures. The defense software base has participated in the first three (though the use of improved programming languages was long delayed by the mandate to write in Ada), but it has not leveraged the last two nearly as much.

Definitions of "agile development" invariably lead to arguments among both advocates and skeptics, but in general the phrase refers to a strategy of rapid, small-scale, incremental development and release of software functionality, driven not by prespecified requirements or specifications but rather by close, iterative interaction with future users of the software being developed. The key features here are as follows:

- **Small**—Features are added in many small increments, rather than a few large blocks/versions/updates.
- **Rapid**—New releases happen on a scale of weeks, not months or years.
- **No fixed requirements**—Users and developers together explore the space of potential features and discover which are the most useful.
- **Interactive**—Users and developers participate as a collaborative partnership, rather than as customer and vendor, with developers in self-organizing teams.

All of these key features pose problems for traditional DoD acquisition. Having many small incremental releases of functionality breaks the logistics system whereby new software releases are coordinated and deployed to far-flung operational units. The absence of fixed formal requirements is antithetic to the Joint Requirements Oversight Council (JROC) mission of specifying formal, validated requirements with threshold levels. It may also cause legal and practical headaches for the writers of requests for proposals and the awarders of contracts, not to mention cost and schedule estimators. The interaction

[4] For our purposes, *improved* simply means more FP of product per SM of effort on average.

between developers and users requires active, ongoing participation of uniformed and civilian personnel who would traditionally never get near the system under development until (perhaps) Operational Test and Evaluation. That ongoing collaboration might last for years.

The other dominant recent development in the commercial world that has generated significant productivity gains is the use of modular open system architectures. Stephen Welby (2014), during his time as Deputy Assistant Secretary of Defense for Systems Engineering, described these as "technical architectures that leverage technical standards to support a modular, loosely coupled and highly cohesive system structure." There are actually two distinct and separately important ideas here: modularity, which is about the way the software's functions are organized into independent composable units, and openness, which is about who can see, modify, publish, or use the code. Not all modular architectures are open; not all open source software is modular. There is a synergy between the two ideas, however—modularity increases the efficiency of individual contributions to the open code base, while openness allows more individuals to contribute.

For our purposes, the key features that drive enhanced productivity are the following:

- Composable software modules that can be combined in many ways to execute more complex functions
- Well-defined, standardized, documented interfaces for these modules
- Universal transparent access to (nearly) all of the source code
- Extensive rights to modify or enhance existing source code
- A large base of independent agents actively engaged in developing/improving the set of modules

Examples of thriving modular open software ecosystems include the Linux operating system, the Apache web hosting platform, the FreeRTOS real-time operating system for embedded systems, the R and Python programming environments, the emacs document editor, and the MySQL relational database. The collaborative nature of the communities of developers working with these tools can lead to enormous total effort—the Linux Foundation estimated in 2008 that the total cost to develop the Fedora 9 distribution of Linux (including the Linux kernel itself) from scratch would have been more than $12 billion (McPherson, Proffitt, & Hale-Evans, 2008). That was nearly a decade of additional development ago.

Modular open architectures enhance productivity through three principal mechanisms: reuse, parallelism, and scrutiny. Modularity allows large parts of the code base to be reused in new applications with little or no modification, greatly reducing development times. It also makes it easier for program managers to decompose complex development projects into weakly-dependent subprojects, so that less work lies on the critical path. Openness, on the other hand, invites large numbers of developers to work on continuous improvements to the ecosystem, so that there is an ever-richer set of existing modules to re-use. This widespread active attention to improving the code in turn results in a higher level of scrutiny—and thus generally lower defect rates—for frequently-used modules in such environments (Brockmeier, 2003). Similarly, software assurance and cybersecurity can be easier for open source software than for proprietary software (Wheeler, 2010).

The openness and transparency of open source ecosystems also provides a welcome indirect mechanism for opening defense software development to the non-cleared workforce. Any defense software that is based on Linux, or written in Python, or implemented using FreeRTOS, is leveraging the efforts of thousands of developers outside the usual defense workforce. In the end, this might be the best argument in favor of modular open source—that it promises not only significant productivity gains for individual programmers, but also the largest available expansion of the defense software workforce.

Recommendations

Thus far, we have presented some plausible guesstimates concerning actual supply and demand, some optimistic yet sobering forecasts, and an enumeration of possible policy options. Given all of that, what should be done?

First, collect data. Study the industrial base; measure the effective demand; measure the maintenance efforts. The forecasts in this paper are built on sparse data from inconsistent sources. An improved update to the Chao (2006) investigation of the state of the defense software industrial base is long overdue, and could replace those credible guesstimates with actionable facts. If we discover that supply has kept pace with demand just fine over the last decade, good. We will have learned something about how the unique defense labor market responds to internal demand surges and competition from the commercial market. If, however, we discover that significant amounts of software maintenance are being deferred, all projects are understaffed, and new programs are executing by stealing from existing programs, then we can sound the alarm.

Second, adopt commercially proven productivity-enhancing acquisition models. In recent decades, the DoD has bet that the boom in commercial software is a rising tide that would lift defense software productivity as well. This turned out not to be true; the needs of the DoD are sufficiently different from those of the commercial world that productivity advances arising in the commercial sector did not necessarily translate to the defense sector. Agile development and large-scale telework are good examples of productivity multipliers in the commercial sector that are not as useful for defense without significant adaptation.

In particular, embrace modular open source software ecosystems. Of the known productivity enhancers, this is the only one that might potentially provide both ongoing rapid productivity growth and an effective expansion of the workforce. Doing this would require substantial regulatory, cultural, organizational, and perhaps legal changes across the defense acquisition enterprise and the defense industrial base. There is also a nonzero risk that such efforts could fail to produce the critical mass of actively engaged developers necessary to realize the benefits of open source ecosystems. Evidence from the commercial world suggests that not only would it be worth the risk, it might be necessary in order to keep up with the pace of technology change and threat evolution. DoD leadership have been pushing in this direction (DoD, 2017), but there is considerable institutional inertia and active resistance to be overcome, both within government and within the industrial base. Furthermore, the early stages of developing such ecosystems might well not look much like progress.

Finally, fund basic productivity research the way the DoD used to do. Without fundamental improvements in software productivity, weapon system capabilities will be limited by the time it takes to develop new software-intensive systems, and that limit may not be very far beyond what is currently being produced. The DoD has little use for highly capable systems that take 25 years to field. In the long run, the key breakthrough will be the automation of software development as a process, so that it is no longer a manual craft labor activity. That vision—autonomous systems writing software from scratch with the dependability required of defense systems—is currently still in the realm of science fiction.

References

Arrington, M. (2010a). Google making extraordinary counteroffers to stop flow of employees to Facebook. Retrieved from https://techcrunch.com/2010/09/01/google-making-extraordinary-counteroffers-to-stop-flow-of-employees-to-facebook/

Arrington, M. (2010b). Google offers staff engineer $3.5 million to turn down Facebook offer. Retrieved from https://techcrunch.com/2010/11/11/google-offers-staff-engineer-3-5-million-to-turn-down-facebook-offer/

Barnow, B. S., Trutko, J., & Piatak, J. S. (2013). How do we know occupational labor shortages exist? *Employment Research, 20*(2), 4–6. doi:10.17848/1075-8445.20(2)-2

Brockmeier, J. (2003). Comparing free and proprietary software defect rates. Retrieved from https://lwn.net/Articles/22623/

Bureau of Labor Statistics (BLS). (n.d.). *Occupational outlook handbook, 2016–17 edition.* Retrieved from https://www.bls.gov/ooh/computer-and-information-technology/software-developers.htm

Bureau of Labor Statistics (BLS). (2017). Occupational employment statistics. Retrieved from https://www.bls.gov/oes/tables.htm

Chao, P. (2006, October). An assessment of the national security software industrial base [Briefing]. Retrieved from https://www.csis.org/analysis/assessment-national-security-software-industrial-base

ClearanceJobs.com. (n.d.) Retrieved from https://www.clearancejobs.com/jobs?keywords=software

ClearanceJobs.com. (2014). In the clear: Compensation decline for security-cleared professionals levels off despite strong headwinds. Retrieved from https://about.clearancejobs.com/hubfs/pdfs/ClearanceJobs_Compensation_Survey_2014.pdf

Department of the Army. *Army Organic Industrial Base Strategic Plan: 2012–2022.* Retrieved from https://www.army.mil/e2/c/downloads/276549.pdf

DoD. (2017, February 23). DoD announces the launch of "Code.mil," an experiment in open source (No. NR-077-17) [Press release]. Retrieved from https://www.defense.gov/News/News-Releases/News-Release-View/Article/1092364/dod-announces-the-launch-of-codemil-an-experiment-in-open-source?source=GovDelivery

Dvorak, D. L. (Ed.) (2009). *NASA study on flight software complexity*. Pasadena, CA: NASA Jet Propulsion Laboratory, California Institute of Technology.

Galorath, Inc. SEER by Galorath. Retrieved from http://galorath.com/software_maintenance_cost

Jones, C. (2000). *Software assessments, benchmarks, and best practices*. Boston, MA: Addison-Wesley.

Jones, C. (2013). Function points as a universal software metric. *ACM SIGSOFT Software Engineering Notes, 38*(4), 1–27. doi:10.1145/2492248.2492268

Kyzer, L. (2017). Clearance salary trends in 2017. Retrieved from https://news.clearancejobs.com/2017/01/19/clearance-salary-trends-2017/

Lauerman, J. (2015). Coding boot camp enrollment soars as students seek tech jobs. Retrieved from https://www.bloomberg.com/news/articles/2015-06-08/coding-boot-camp-enrollment-soars-as-students-seek-tech-jobs

Longstreet, D. (2001). Software productivity since 1970. *Function point training and analysis manual.* Retrieved from http://www.softwaremetrics.com/Articles/history.htm

Lucero, D. (2009, April). *Software sustainment challenges in defense acquisition*. Presented at the AIAA Infotech@Aerospace Conference, Seattle, WA. doi:10.2514/6.2009-1816

McPherson, A., Proffitt, B., & Hale-Evans, R. (2008). Estimating the total cost of a Linux distribution. Retrieved from https://www.linux.com
/publications/estimating-total-cost-linux-distribution

National Research Council, Committee for Advancing Software-Intensive Systems Producibility. (2010). *Critical code: Software producibility for defense*. Washington, DC: National Academies Press. doi:10.17226/12979

System Architecture Virtual Integration (SAVI). Retrieved from http://savi.avsi.aero/about-savi/savi-motivation/exponential-system-complexity/

Tate, D. M. (2016). *Software development may drive future acquisition cycle times* (IDA Research Insights NS D-8053). Retrieved from https://www.ida.org/idamedia/Corporate/Files/Publications/Insights/CARD-D-8053-Tate.ashx

Welby, S. P. (2014, October). *Modular open systems architecture in DoD acquisition*. Presented at the 17th Annual NDIA Systems Engineering Conference, Springfield, VA.

Wheeler, D. A. (2010, April). *Open source software (OSS/FLOSS) and security*. Presented at the International Workshop on Free/Open Source Software Technologies, Riyadh, Saudi Arabia. Retrieved from https://www.dwheeler.com/essays/oss_security_saudi_arabia.pdf

Acknowledgments

This work was performed under the IDA Central Research Program, supported by Department of Defense contract HQ0034-14-D-0001.

Transformation of Test and Evaluation: The Natural Consequences of Model-Based Engineering and Modular Open Systems Architecture

Nickolas H. Guertin—PE, is a Senior Systems Engineer at Systems Planning and Analysis, Inc. He is a recently retired Navy Civilian, where he held a variety of technical and business leadership positions, including Director for Rapid Prototyping and Strategic Planning at NSWC Carderock and Director for Transformation in the Office of the Deputy Assistant Secretary of the Navy for Research, Development, Test, and Evaluation. He holds a BS in Mechanical Engineering from the University of Washington and an MBA from Bryant University. He is a registered Professional Engineering and is DAWIA certified in Program Management and Engineering. Guertin is also a retired Navy Reserve Engineering Duty Officer.

CAPT Gordon Hunt, USN—is a Naval Reserve Engineering Duty Officer and supports PEO IWS and NAVSEA on naval combat system software solutions and design principles. His civilian employment (Skayl) focus is on system of systems integration and semantic data architectures, enabling increased systems flexibility, interoperability, and cyber security. He holds an MS in Aeronautical Engineering from Stanford University and a BS in Aerospace Engineering from Purdue University. He also serves in industry leadership as an appointed member for the Advisory Board of the Open Group's Future Airborne Capability Environment (FACE) architecture consortium.

Abstract

This paper examines the technologies and architecture patterns that are transforming software-intensive systems and the Internet of Things (IoT) that are currently being designed and implemented. The use of these practices should create an ensuing transformational shift in the relationships between the Test and Evaluation (T&E), development, and operational communities.

Based on the findings of this research, a set of practices for a coordinated set of hardware, software, functional, and data architecture patterns and testing strategies is presented. This paper will show how these need to be applied via a data architecture that defines the declared test points between modular components in software intensive systems. This will support affordable and rapid integration of innovation through a business model that uses small-scale component replacement. This research ends with an assertion that, when the right architectural elements are standardized, regular incremental improvement is both affordable and effectively applied throughout system development.

Introduction

This paper proposes a new path toward a robust and affordable approach for product development to achieve the fundamental purposes of T&E—to validate and verify the acquisition of excellent military capability. The architecture itself, not just the content, should also be testable to its own set of requirements. As such, there needs to be a set of practices that can directly test such architecture characteristics of flexibility, scalability, interoperability, and so forth, prior to making major investments in detailed development. Then, when the content of the components that make up the system are filled out, the test and evaluation process can validate and verify that the content is following the constraints of the architecture. In this way, when the full system is completed, the program is not at the beginning of traditional T&E, but at the end of the product development and test process, and can quickly transition to fielding. Synergistically, when open and functional architecture steps are followed, some of the smaller testable chucks of software prior to the "full system" being created, thus expanding opportunities for broader enterprise value of strategic reuse.

DoD procurement is changing, driven by a combination of the national strategic imperative to much more rapidly address the needs of the warfighter who is facing an asymmetric enemy, who is also able to access the fundamental underlying building blocks of capability derived from globally available commercial technology (National Defense Authorization Act, 2017). As a result, defense procurement needs to become nimble, deriving new mission capability through flexible and rapid integration of capability modules instead of classically procured standalone systems (Richardson, 2016).

A natural extension of these assertions is that fielding of new capability must also be made more fluidly, frequently, and in smaller increments than the large-scale major systems deployments typical of classic Program of Record approaches that follow a design/build/field/sustain/dispose life cycle as shown in Figure 1. The environment must adjust to a different deployment model for capability where new features and performance capabilities can be delivered when they are needed in the field.

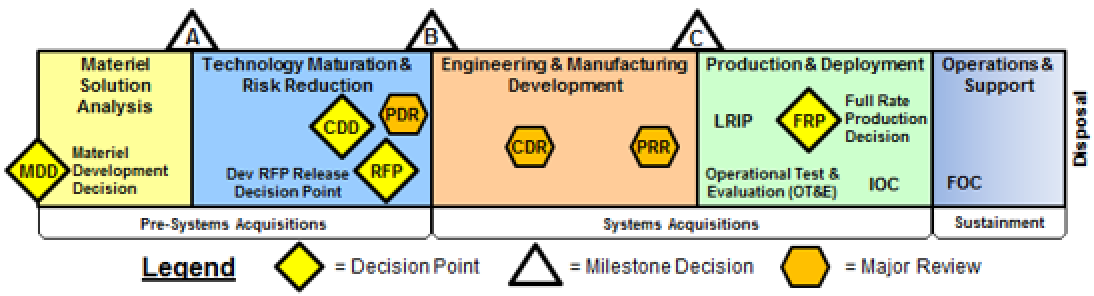


Figure 1. Defense Acquisition Framework

The new model changes the design/build/field/sustainment approach of discrete and separate phases into one of continuous engineering and deployment. In order to achieve this, new increments of capability, acquired from a wide range of offerors, must be able to be affordably tested and fielded within days or weeks. Both the testing community and the acquisition environment must have strong evidence and buy-in that such an approach can be risk-prudently performed. This will be a difficult change in culture as both of these communities are steeped in the natural cadence of the defense acquisition framework, which can require years to move from characterizing a problem to having a fully fielded system-specific capability.

The Changing Environment

Energetics and life-costing decisions hang in the balance of military products. The rigor of testing and certification required for warfighting capability is very different from standard industry practices for consumer products, such as testing an incremental release of a popular mobile application. As such, software for military warfighting systems, regardless of its origin, must be governed and implemented with rigor. This is made more urgent by the growing and persistent cyber threat to software-intensive systems. Consider the following observations on current DoD software architecture practices:

- The System's Engineering "V" diagram is being eclipsed to today's by new forms of robust model-based systems engineering and systems-of-systems design environments and tools (Micouin, 2014). In addition, system performance requirements must adjust as the technology matures and the warfighting problem space changes. Architectures and test capabilities are needed to readdress the duration of the continuous engineering, testing, and deployment phases.

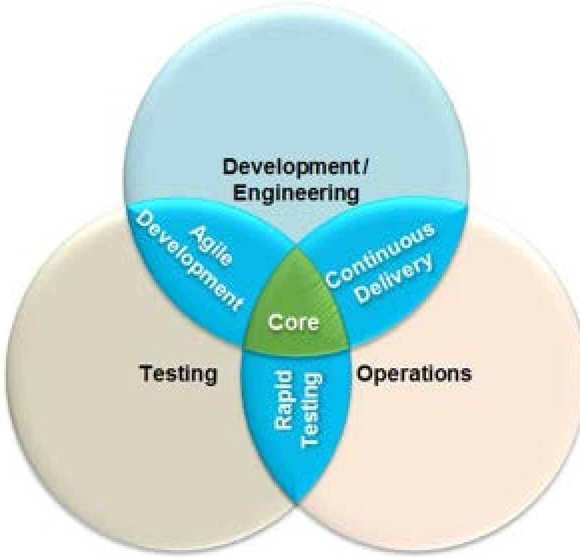


Figure 2. Continuous Engineering
(Rahman, 2014).

- Modern cyber-physical software development practices are using decomposition of capabilities into smaller, individually competed functions (Guertin, Schmidt, & Sweeney, 2015). These functions also need to come built with accessibility to internal software “test-points” and conformant external interfaces. Both of these will be necessary to support automation of tests.
- Today’s delivered capabilities span multiple systems, programs, and services (Jamshidi, 2008). The overarching integrated capabilities are composed of orchestrated behaviors forced through an array of architectures, deployed on different hardware, using different internal interfaces with a multitude of different data representations.
- Advances in Model Based Engineering (MBE) now enable the acquisition community to explicitly address integration complexity for definitions of system software specifications (DoD, 2017). This action will fundamentally address the one issue that simultaneously decreases system costs and reduces time to deployment.
- The Defense Department Services are revolutionizing software development with researched, tested, and validated Open Architecture approaches (Guertin & VanBenthem, 2016). However, successfully delivering its full value to the warfighter requires the entire DoD procurement cycle to fully embrace its potential and to deliver on its value.
- The T&E community also develops large, complex software intensive systems to support testing (Deputy Assistant Secretary of Defense [DT&E], 2016). Those MBE practices need to be aligned with the associated MBE efforts being used in the acquisition community. The T&E community has the ability to develop a rich set of cohesive testing infrastructure and tools while still preserving their independent role.

Testing assembled systems and systems-of-systems (SoS) has been long, complex, and expensive. As such, the effort associated with fielding adequately tested products is rapidly increasing (Deputy Secretary of Defense for Systems Engineering, n.d.). Automation of testing activities is necessary to address these activities to improve test breadth and penetration, while simultaneously increasing speed of delivery and reducing the overall test burden (Elfriede, Rashka, & Paul, 1999). This will similarly require robust testing frameworks that are not intimately tied to the product being tested. To align the product development and system test domains, the full panoply of complexity must be addressed, such as: internal component functional testing, system integration, and cross-system behavioral testing of software-intensive systems.

This creates an opportunity for alignment of the MBE efforts across the acquisition workforce. The exploding complexity can be managed through a greater emphasis on defining the data artifacts of the modules of the systems under test, while also enabling their extensibility for re-use. This includes tools that can specially address the complexity of testing software intensive systems, a market-place of T&E products and test artifacts, and a third-party marketplace of innovative products to support the T&E workforce.

Design and testing of interfaces are also going through fundamental changes in both approaches and results. The classic approach of an Interface Control Document can be replaced by using a combination of MBE and a supported data model. These together fully define what data moves across the system and how data is used internally to a module to support functional performance. The consistent and testable MBE processes and software architectures can then be used to provide “managed” automation of testing and interoperability. This results in testing products that are open to support integration and are readily reusable across programs.

Defining a “Testable” Architecture

The T&E community can test components and software early and often by first decomposing the criteria for the fundamental building blocks of software intensive systems. An analogy is beneficial to set the stage—the most accessible and reliable one is your house.

To better understand the relationships of enterprise design, consider the comparative example of how communities build out their towns. Figure 2 depicts the relationship of the enterprise architecture to the community’s master plan. The enterprise architecture is the first tier of a multi-level design process. Both the large-scale plan and the individual house plan represent a forward-looking vision of the eventual community or product-line implementation. Both the developer and the inspector are governed by regulatory practices and architectural patterns and styles, and they must be responsive to future market and business-driven factors. In short, building codes define the architecture rules.

At the highest level in the building architecture analogy, business and community leaders determine what they want their town to look like and what infrastructure requirements they might need. Roads, utilities, and capacities are examples of the highest level requirements of a community. Likewise, an enterprise product line for a defense system constructs a set of rules to facilitate systems domain business requirements, such as portability, reuse, interoperability, speed, and scalability. These requirements are then translated into attributes that the resulting architecture must possess.

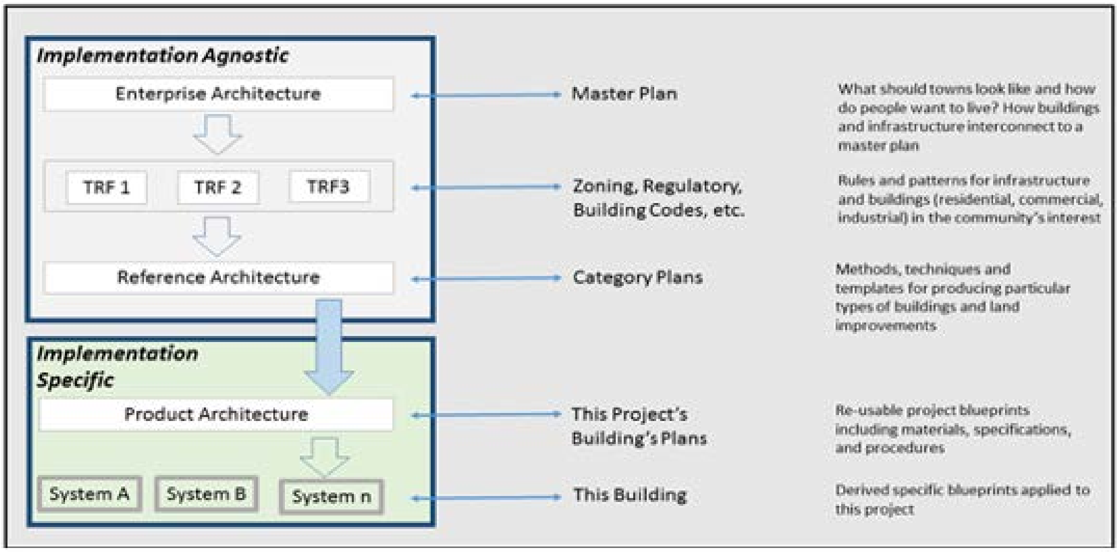


Figure 3. Architecture via Building Codes Example

When a house is built, it must conform to these overarching requirements. The rules of construction are set for things like framing, electrical, plumbing, insulation, internal and external finishes, and so forth. Building/zoning codes and other constraints, like homeowners association agreements, can also assert controls on how the building interacts with the rest of the community.

The home builder does not set the building codes or the inspection and test methods. The inspectors have a basis for evaluating creative alternative implementations while preserving safety for the individual and value to the community. These codes and building rules must be structured to be loosely coupled and have limited impact on other rules. However, when new construction methods, modern materials, or new aesthetics are presented, the test criteria and test methods must react dynamically and evolve.

The physical nature of a building forces us to take a step-wise approach to inspection, such as the overall design is inspected before construction begins, the foundation is inspected before the structure is built and before electrical and plumbing is installed, etc. Software intensive systems could be approached in a similar fashion. As such, the developers and evaluators would establish a partnership for setting the building codes and creating the criteria for testing and inspection.

Establishing the Categories of Architecture

The T&E community must be involved in defining the "building codes" for today's complex software-intensive systems to establish the governing principles for building, implementing, and eventually testing software-intensive System-of-Systems (SoS).

The scale of the problem is growing faster than existing methods can account for. Consider that a modern automobile can have 10 million lines of "mission critical" code ("How Many Millions," 2017), and even the operating systems can easily top 25–30 million lines of code. Similarly, modern weapons systems can each have in the neighborhood of 10–20 million lines of code. Additional software is needed to integrate and share data between all platforms, sensors, and weapons to make these complex systems perform together. In addition, the very act of testing these software-intensive systems is creating significant amounts of software as well. As a result of this complexity, one or more of the following challenges are often observed when testing and integrating large software-intensive systems:

- Integration patterns limit flexibility for incorporating new capabilities.
- System do not scale in size or diversity and have less capability than required.
- External or key interfaces do not work as specified or designed.
- Functionality is reduced from design specifications.
- Interface documentation is insufficient to effectively test system boundaries.
- Traditional Interface Control Document (ICD) specifications evolve too slowly to accommodate evolving or novel capabilities.
- The combinatorial challenge of testing every interface leads to untested interactions (Kuhn, Kacker, & Lei, 2013).

Analysis of these failure mechanisms often indicates that unplanned dynamic behavior exists among the key system elements (Capilla et al., 2014). Since the characteristic of this failure mode is unanticipated and systemic, it is unlikely that traditional analysis and engineering judgment will provide a robust and enduring solution. The situation will deteriorate as systems get more distributed, complex, and interdependent. The commercial industry is working to address the challenge of an estimated 20 billion connected devices by 2020 (Hosain, 2016). This revolution is happening over time to build and connect these devices, including finding agreements on how they all connect together. The solution to this complexity must be foundational and a fundamental aspect of architecture.

In some isolated cases, the idea of a testable architecture has been realized. For example, Architecture Analysis and Design Language (AADL) was created for the specification, analysis, automated integration and code generation of real-time performance-critical distributed computer systems (Architecture Analysis and Design Language, n.d.). AADL provides additional model-based engineering mechanisms to test an architecture prior to full product development. Outcomes of embracing testable architectures include the following:

- Early detection and debugging of unanticipated dynamic behavior
- Exercise key interfaces early in the development phase
- Enable benchmarking of key function prototypes in the system environment to provide visibility into unanticipated dependencies
- Provide for tools that can test the architecture separate from function enabling repeatable testing of early development products in a system environment
- Analysis and comparison of design alternatives in a system environment
- Discovery of emergent net-benefit capabilities that can be realized when integrating large systems and system-of systems.

The following assertions were derived from the experience of the authors and validated through research:

- Integrating systems in predictable and testable increments reduces risk and rework.
- Defining the rules and the architectural elements that enable this behavior further improves outcomes.

- Analysis of the different documents, standards, processes, models, software libraries, and physical components led to a definition for how to implement a system.

These core concepts were then distilled and mapped to a set of architectural elements from which systems can be informed, specified, designed, and implemented (Allport, Hunt, & Revill, 2016).

Table 1 presents a simplified view of that analysis, where the input in the left column is the software, interface, or hardware specifications that the acquisition community currently leverages in execution and implementation of cyber-physical systems. Those were grouped and identified the input's core architectural tenets to realize the reference architecture categories. As additional input, consideration was given to current architectural tools, standards, and best-practices to ensure that architectural content in an identified category could be captured and documented, and most importantly, tested.

Table 1. Core Architectural Elements

Cyber-Physical System Concepts Leveraged in Execution & Implementation	Core Architectural Tenets	Reference Architecture Category
Interface Control Documents (ICDs) Interoperability Profiles (IOPs) Tactical Data Link Specifications (TDL)	Interfaces Messages	Functional Architecture
CPU & Hardware Selections Network & Communication Fabrics Hardware Input/Output (IO)	Deployment	Hardware Architecture
ICD Documentation Configuration Files Intrinsic Knowledge of Meaning	Knowledge Information	Data Architecture
Infrastructure Software Operating Systems (OS) Middleware Libraries Software Development Kits (SDK) Model-Based Engineering Code Utilities	Applications Infrastructure	Software Architecture
Commercial Standards Defacto Standards DoD Specifications & Requirements	Standards	Functional Architecture
Acquisition Process Contracting Actions Requirements & Specifications	Business Model	Governance

The reference architecture categories serve as the building code categories for specifying, designing, and implementing systems and testable elements in the architecture. Each of the identified reference architecture categories are defined as follows:

- Functional/System Architecture—This architectural segment is closely tied to the business goals of the system and includes statements about what a device or service "does," what it "provides," and what it "needs." Testable KPPs are usually defined against the functions and are implicitly coupled to the implementation requirements. Traditional ICDs document and define messages and interface syntax as aspects of an interoperability requirement on a unit of function. These often implicitly couple a function's deployment and current intended use into its specification. Various model-based engineering tools and standards exist for documenting interfaces of a system. The challenge is to ensure that the documentation and design clearly decouples software, from data, from function of an interface specification.
- Software Architecture—This architectural segment focuses more on how a function should be implemented in code and logic. It covers how the software infrastructure, computational support interfaces, operating systems, middleware technologies (Hohpe, 2004), and display technologies are used and integrated. A software architecture defines the boundaries between components of functionality, the granularity of those components, how those components communicate, and how the resulting software is deployed and managed. Key interfaces are identified, and mechanisms to test and decouple the interfaces are often elements of a software architecture. The challenge is that software architectures are not crafted as enduring designs and many times end up as a *defacto* system architecture coupling one system's implementation specifics to every other software service in the system.
- Data Architecture—This architectural segment is focused on documenting the content *and* meaning of data. Data is not just what is exchanged between functions and comprises more than the messages. Levels of interoperability (Tolk & Muguira, 2003) define not only the structure (syntax) of the data but also the context and behavior (semantics) as well. Traditional documentation of data has captured syntax, but semantic and meaning of the data is implicit when considering model-based engineering practices. Recent standardization efforts and activities have clearly delineated the data architectural properties necessary for a reference architecture (The Open Group, n.d.). This includes the ability to identify the syntax (structure of interfaces and messages), the traceable context (semantics of the data itself), and the behavior of the data as presented through the documented interfaces. The challenge in data architecture is decoupling the documentation of the meaning from the presentation and dissemination of messages. Often what is actually exchanged carries additional meaning and understanding that is not overtly stated or captured. Not capturing this meaning in MBE formats make testing at scale certain to generate integration errors and is the source for many of the non-desirable emergent behaviors.
- Hardware Architecture—This architectural segment covers specifications on physical computer hardware, network fabric and I/O signaling mechanisms, hardware mounting, power and handshaking protocols, connectors' wiring specification, and the like. The T&E community has a long history of

successfully and independently testing hardware. The challenge addressed in architecture is to test and ensure that software, functions, and data are sufficiently decoupled. While it is sometimes advantageous to directly talk to hardware for performance, the T&E community needs testable architectural mechanisms to isolate and decouple software and function from the hardware.

- Governance—This is a critical component in a reference architecture and details where and how the various levels of the architecture will be assembled, deployed, evaluated, and tested for conformance. The architectural categories that need to be put together for testing are as important as the testing architectural products of the individual categories.

The relationship between these architectural segments provides additional testable architectural attributes. When coupled with a functional architecture, the data architecture ensures the information flowing across to peer-level modules will be correctly interpreted when new elements are added. The software and data architecture boundary ensures that information bringing exchange within software libraries is fully documented and understood. The software and hardware architecture boundary establishes the required decoupling and interface abstraction required for portability and extensibility of the implemented functions. These boundaries are especially critical when components that require interaction are crafted by new suppliers or third-parties.

Tying the architecture categories together with a data architecture in this way reduces program risk by easing integration of replacement or new capabilities by adding clearly documented semantics and meaning—something that is lacking in today's MBE tooling.

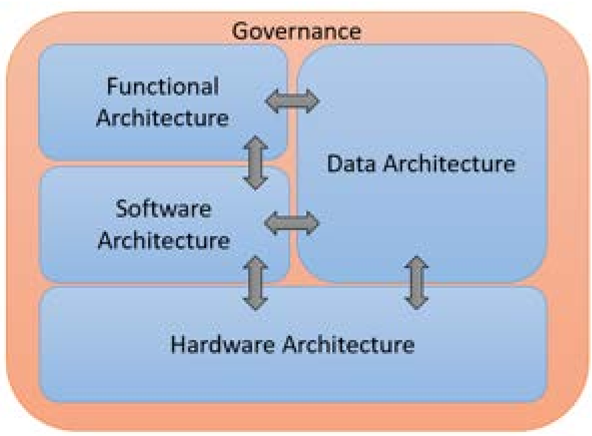


Figure 4. Data Architecture Prominence in System Architectures

These formalized concepts will result in individually testable architectural rules, testable relationships between the rules, designs that can be evaluated against the rules prior to implementation, and final products that are testable throughout the development life cycle. In order to accomplish this, standard arrangements of standards are needed to build and group the governing set of rules. A growing and powerful practice for achieving this is to use Technical Reference Frameworks (TRF; Schmidt, 2016).

Characteristics of a Technical Reference Framework

TRFs define implementation-agonistic design environments and patterns that establish a common set of practices for use in a specified context. In effect, a TRF is a standard for how to use a set of standards to achieve a class of designs. Program managers and their architecture teams can choose TRFs to apply their product requirements against enterprise business drivers, with the goal of creating reusable components and to establish opportunities for any practitioner that can access the environment and add value.

A minimum of three TRFs are needed to craft the full range of military mission systems and for developing reference architectures are shown in Figure 5 (Lethart et al., 2016). Note that the TRFs overlap and transition across the dimensions of criticality and scale. Product requirements may dictate the use of more than one TRF in the development of the reference architecture. For example, a system may require TRF1 for control of a vehicle, TFR2 for command functions, and TRF3 for data analytics of sensor information. TRFs are not aligned with products and systems platforms, but rather the physics-based drivers that guide and constrain how systems get implemented and connected. The three TRFs are summarized as follows:

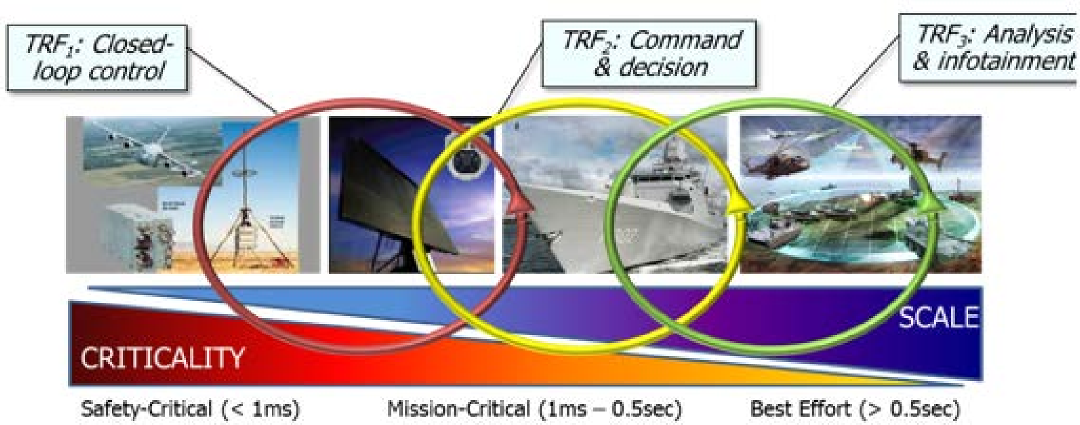


Figure 5. Technical Reference Frameworks and Time Domains

Safety Critical (TRF1): This TRF addresses the most critical requirements for the safe and continuous operation of the system or platform, as well as the most demanding design requirements such as personnel or weapon safety. Safety critical requirements are the ones that must take precedence if there is a conflict with other technical aspects of product. Testing these products require high degrees of timing precision and is often coupled to real-world dynamics. Implicit in the design of high availability systems/functions are forms of internal redundancy, unit duplication, direct control and polling of separate solutions, and dedicated allocation and management of resources. However, in the context of TRFs, those patterns remain implementation-agnostic. When applied, the segments of the system designed to TRF1 will meet the highest level of criticality.

Mission-Critical (TRF2): This TRF is applied to functions that comprise the mission capabilities of the platform. Timing and scale are the prime drivers within TRF2. The purpose of TRF2 is to apply modular, data-centric, loosely coupled solutions (e.g., using inversion of control patterns) to create architectural elements that satisfy performance requirements, with stringent end-to-end timing and reliability quality attributes forming key design decisions. Subordinate requirements, such as scale, regulatory compliance, and security are applied in a recursive fashion until all requirements are met.

When applied, TRF2 will manage performance of designs that are highly time sensitive, but not safety critical.

Analysis and Support (TRF3): This TRF is applied to portions of the design that has low criticality, i.e., they may not need to operate on strict time deadlines and or be hardened to survive in harsh environments. Like TRF1 and TRF2, the elements of TRF3 must adhere to the enterprise architecture model as quality attributes of an integrated system or system of systems reference architecture are created. Program managers generally would use this TRF for products that address capabilities associated with analysis, support, and infotainment applications. This analysis is guided by TRF3 patterns, which often involve virtualization, containers, and resource pooling/sharing.

The architecture team, a collaboration between the program office and the T&E community, should evaluate system requirements and assign the appropriate TRF(s) to guide the development and of their reference architecture. The reference architecture then guides and supports continuous testing throughout the development life cycle.

Testing a Design Using a Reference Architecture

Continuing with the building code analogy—an electrical inspector need only worry about the electrical concerns of a project. Whether the structure is business or residential, there are existing guidelines which dictate practices for wire gauge and placement of electrical outlets. Inspections can be performed in phases as construction proceeds, and if an inspection fails at any point, the errors must be remedied before work can proceed. Additionally, there are cross-cutting specifications when electrical passes through framing or is near plumbing and water fixtures.

Instead of testing to a common implementation specification, the architectural-rigorous approach tests against a set of design tenets defined in the reference. Testing allows determination of these independent design elements. The disparate components of the product can then be properly integrated. For example, a builder does not have to wait until the house is fully built to buy faucets. There are specifications that govern the interface between the faucet and the counter and allow one, two, or three holes at various sizes in the counter. These same principles apply with TRFs. Flexibility is preserved with the use of the interfaces between and in the reference architecture categories, without resorting to being forced into reusing legacy implementations that add fragility to the end product.

Each previously introduced aspect of the overall reference architecture, software, hardware, functional, and data, has its own set of test points, tools, and MBE-based documentation practices. This needs to be performed as a carefully considered deconfliction of the related standards and specifications.

The location of these test points establishes considerations for the automation afforded by MBE-based approaches, to include following:

- The T&E Community being engaged in defining and *maintaining* the architectural elements enables testing early and often.
- The product designs can be tested against the architecture well before they are integrated into the system, e.g., testing the faucet design without knowing where it will be installed.
- The ability to test at all levels of the design is established in the context of the architectural elements, e.g., testing whether a home has enough bathrooms doesn’t require detailed design of the entire house.

- Creativity in the solution space is preserved by testing a design against architecture principles, vice against a specified implementation.

In short, each reference architecture category is required to be documented with its set of building codes. This defines architecturally where the test points are well before a Program of Record starts. Certainly more can be covered and detailed, but for the purpose of this paper, a set of key testable architectural elements and patters from each reference architecture category are highlighted.

Software Architecture

Many books have been written about good software design. They agree on the fundamentals, but then diverge in different directions to add specialized guidance that may be less acceptable depending on what is being implemented. While a TRF is a selection of design-appropriate constraints for the system, there are a few common elements of software architecture that the majority of systems leverage. Highlights include the following:

> Design for Orthogonality: Functional units of software are built to perform a specific function and can be swapped with other implementations that perform that same function (with new or improved features). When functional units do one thing, they are constrained to not be doing something else. Despite the obvious nature of that statement, behavior leakage is inadvertently built into software all the time. Products designed to do a single thing can be tested and do not require the rest of the system for a valid result. Occasionally, these tests can be performed in a simple test harness. If the software design is sufficiently granular and the software decoupled, this can allow for early testing.

> Minimize Coupling—Interface-Based Designs: Software changes. Implementations change. Designs that focus on a consistent interface are far more resilient to change. This principle is very similar to the notion of orthogonality, but refers to how the software is built rather than isolating functionality. The amount of coupling is an engineering tradeoff that should be most addressed overtly at the end points of decision strings or edges. The location of those edges are a function of the TRF selected and can be at a software library, system executable, or entire virtual operating system.

> Test-Driven Design: Another popular methodology is test-driven design. Before any implementation software is developed, the test cases for the corresponding requirement are developed. Then, the code is written to pass the test. This ensures that, at the very least, the requirement is met. The challenge is that designs of tests based on current understanding of the product and how it is expected to function, and those implicit assumptions, are part of the tests and developed software—for example, deciding that the test for a bathroom counter will test for three holes. Unfortunately, the right level of specification is driven by where and how integration flexibility is required, hence the need for the TRF to define the rule sets.

It is not necessary to implement a full test-driven design process in order to benefit from these principles. The value added is that software is tested against an architecture as it is developed, ensuring the design and deployment requirements of the TRF are met. Current MBE-based software engineering tooling provides requirement traceability, test artifact generation, machine parseable documentation of deployment and system topologies, code coverage, other analysis tools, and more. Each serves a function, but without the traceability to a reference architecture "rule" there isn't a consistent application of their utility.

Hardware Architecture

The T&E community has a long and successful history testing hardware. While this paper is focused on the current software complexity challenges facing today's integrators, there are good examples and lessons from the hardware perspective. The JTAG (IEEE, 2009) standard is just such an example. It provides the physical test point specification as well as the data and signaling IO that the interface supports. The JTAG community didn't invent new technology, but rather assembled a standard of standards (serial communication, power, connectors, etc.) to support their use-cases and defined data in the hardware test domain.

There are parallels in the JTAG standard and what this paper is proposing. Clearly defined separations between the hardware, software, and functional architectures provided the flexibility and endurance of the standard over the past 20 years. How the data is exchanged and packaged over the interface is well understood, but what the data means from a particular device is often documented separately. This highlights what has traditionally been the gap in architecture specifications. The syntax can document a system's interface on MBE-based tooling, but little is done other than human-interpreted prose to document the semantics. This makes every integration a process of discovery of the meaning of the data. Fundamentally, it is the data architecture that has been the missing testable piece.

Data Architecture

Data architecture for interfaces is the newcomer to the conversation. Data isn't new, and database administrators have had an architecture to describe and document data's meaning for quite some time. But about data in motion? In the past, developers have been allowed to simply "create a new message" to communicate system information and state. This approach is sufficient so long as systems remain relatively small and not connected to too many external systems.

Integrating with external systems requires documentation. It requires understanding of the data's format as well as the data's meaning. To date, most development teams have relied on paper-based ICDs and some MBE-based representations of the ICD. While these documents are fine for capturing syntax (the structure: units and data type), they fall short on capturing the semantics, or, what the attributes actually mean. By adopting a rigorous approach to data architecture, the syntactic and semantic rigor in a machine readable and machine-understandable data model can be captured. It is the machine-understandable part that enables MBE-based methods for automating testing data and the interfaces that exchange it. Much like measuring and testing a database against the normal forms (Kent, 1983), testing can instead use the semantic documentation of interfaces for completeness and machine-based utility.

A powerful new integration tool is created once the exercise of capturing this information has been performed. The meaning of data is also captured when the semantics of data have been documented in a machine-understandable format. This means that a computer is able to mathematically process equivalence relationships with absolute certainty, not using stochastic processes. *A priori* integration and analysis of the data exchanges can be performed when starting the design process armed with data architectures. These systems then can be related to each other and analytically determine the overlap/gap of the integration using machine-understandable documentation (Hunt & Allport, 2016).

Data architectures test the structure of the documentation and make explicit the meaning of the content. The definition of the meaning (semantics) of the data in the interface in MBE-based formats allows testing of the interfaces with a defined syntax of their content. A data architecture supports a rigorous functional architecture.

Functional Architecture

Finally, the functional architecture needs to be explored. This reference architectural category is interesting in that many aspects of it get coupled inadvertently with other architectural specifications. For example, the interface is designed to accommodate the hardware's limited bandwidth, address implications of limited compute resources, enforce a singular use of the data, or bake in signaling protocols that dictate implementation patterns. The first decoupling aspect of a functional architecture is to separate interfaces from messages. Often these are implemented as the same thing, which results in significant coupling in order to have an architecture that manages and treats the interfaces and messages as separate testable specifications. Further decoupling includes the functional decomposition of the system itself. This decomposition details each component's role in a system. In order to test the decoupling of the roles in a functional architecture, there are several approaches and techniques (a TRF will have to specify what applies for its domain) that can be leveraged. This is certainly not an exhaustive list, and the interaction with the other architectural would need to be defined.

- Functional Flow Block Diagram. The diagrams define the step-by-step flow of the logical order of execution in the system. In UML, these types of diagrams can be manifested as sequence diagrams. Other standards exist for documenting this implementation detail.
- N^2 Chart. These are used in software-intensive system to calculate the coupling between the inputs and outputs of the identified functions (Hitchins, 2003). A decoupled functional architecture limits the dependencies across components in the system and makes it easier to predict impacts of updates to a component in the system.
- Structured Analysis and Design. This methodology can be leveraged to describe systems as a series of functions, with identified inputs, outputs, and supporting data and mechanisms for the function's action. Again, it provides a way to quantify (test) the robustness of a design.

Testable Architecture Summary

In summary, the data architecture provides the binding and traceability between the different architectural segments. The architectural specifications needed to build a design with increasing levels of specificity can link the implementation all the way back to the architectural goals inherent in the applied TRFs. Each level adds detail about the location of the interface test points, the data and its meaning that is exchange over the interface, and methods and implementation patterns that minimize coupling of software. In the end, software is just a specification that is compiled into many very specific instructions for a processor. The specification of a system and its interfaces can be treated with the same level of testable rigor as a software program as long as the right architecture is in place. The assembled systems can then be tested with compliable automation versus human-powered actions.

The Consequence of Testable Architecture

System and system-of-system testability is enabled by an open architecture. An open architecture is achieved when the rules-of-construction are clearly documented, deconflicted, and documentable in model-based engineering processes per the selected TRF specification (Figure 6). When a program adopts an architecture specified to this level of rigor, the infrastructure itself becomes an explicit and a separate component of that system. The open infrastructure changes the nature of integration as well the relationship between the T&E and development communities.

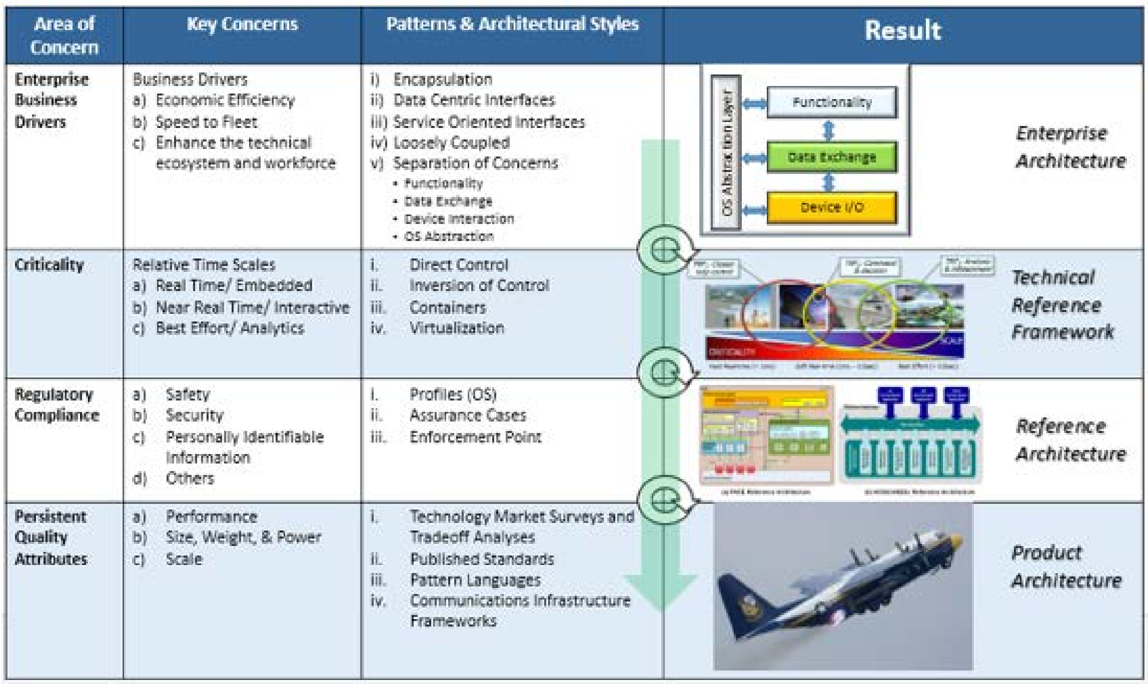


Figure 6. Realizing a Product Architecture Through a Series of Testable Steps

The infrastructure can be maintained and provided separately from the platform functional software. Many companies could build modules or systems to the infrastructure and any qualified vendor could then integrate a component and have it function in the system. As an example, the same plumber that built the home is not needed to replace a faucet. Your home is in fact a testable, open architecture, from the design, through the construction, and the following years on maintenance and updates. An open architecture requires an open infrastructure, an open acquisition business model, a technical and operational roadmap, and an organization that can support and maintain these items.

An open infrastructure has three primary characteristics: The first characteristic of an open infrastructure is that it has an open data model that is rigorously defined, described, and fully discoverable. The data model must be completely published, and based upon an abstraction that is broad enough to define the full domain of the system. An abstract data model enables a model to achieve the full breadth of possible implementations, while also defining repeatable interoperable mappings between these possible implementations.

The second characteristic of an open infrastructure is that it is based on open standards and is flexible. Open standards do not limit differentiation, innovation, or competition and they ensure a commodity infrastructure. Flexibility is as important as open standards because there is no one technology or application programming interface (API) that is sufficient for the range of behaviors necessary for a complete, enterprise-level infrastructure. This flexibility is achieved through the use of architectural patterns at the service and interface model. The patterns specify expected behavior while not over-constraining the communications design and build of the infrastructure architecture.

The third characteristic of open infrastructure is that the infrastructure can be developed, acquired, and maintained independently of the functionality of the system. This enables functionality to be acquired independently of the system infrastructure and ensures that capabilities are delivered without subsystem and special-purpose infrastructure dependencies.

Several organizations are achieving these ends. For example, the Army has been doing so for its next generation ground control software (Bellamy, 2014). By defining their software architecture, the decoupling between the hardware and software architectures, and elements of the data architecture, they are laying the groundwork for a generation of incrementally testable components and software.

Cultural and Organizational Impacts

The building code analogy is a powerful example of a robust design and production market where creativity is highly prized, while overarching public good is managed. Architects, standards bodies, contractors, inspectors, and consumers work together to ensure new and exciting products are available to the customer.

Similar to the relationship between the inspectors and the builders, the T&E community must preserve its arms-length relationship with the development community. In this way, they can ensure that products are built with the requisite capability and inherent flexibility to grow over time. To facilitate that shift, the test tools and capabilities must be grounded in making sure that systems have the right architectural features as well as making sure that the unique military capability is delivered.

This change in relationship places the T&E role much earlier in the development cycle and in partnership with the development and operations efforts. They must be a part of setting and evolving the standards going forward and ensuring that the test products address fundamental architectural principles, versus purely on operational capability.

Conclusion

The historic path of product development will lead to accelerating growth in complexity with unsustainable increases in development and test time and cost. An approach to crafting software intensive products needs to change to address complexity of capability while simplifying the way those products get built, tested, and fielded.

TRFs can be used to establish the “building codes,” or architectural patterns, for the product based on what the design needs to accomplish. As a team, the operators, program manager, development engineers, and test & evaluation experts can select the rules of construction during the early stages of product definition. This becomes the reference architecture and the T&E community sets the stage for how the product will get tested before it gets built—establishing the testing life cycle for the product. Testing starts with validating, early in the product development life cycle, that the architecture of the design will support the intended performance of the requirements. T&E engages early and stays involved throughout development, finding problems as early in the cycle as possible, correcting them where it is efficient and effective, driving down cost, lowering risk, and increasing robustness.

A natural consequence of these practices is to end up with a new and separate component of the design: the infrastructure. Furthermore, by infusing that infrastructure with configurable variation points, a product line architecture is created that can be used to quickly instantiate alternative downstream implementations—a critical enabler for enterprise reuse. In addition, by using TRFs that are widely practiced, any qualified vendor can create capabilities that can be added to the product line. Lastly, the tools to the trade of integration and test are known and practicable by that same community of practitioners, such that the role of integration or testing can be risk-prudently performed by alternative vendors.

References

Allport, C., Hunt, G., & Revill, G. (2016). Rethinking system integration*. Electronic Engineering Journal*. Retrieved from http://www.eejournal.com/archives/articles/20160609-interoperability

Architecture Analysis and Design Language. (n.d.) What is AADL? Retrieved from http://www.aadl.info/aadl/currentsite/

Bellamy, W., III (2014, September 8). New software framework coming to Army UAS ground stations. Retrieved from http://www.aviationtoday.com/2014/09/08/new-software-framework-coming-to-army-uas-ground-stations/

Capilla, R., Bosch, J., Trinidad, P., Ruiz-Cortés, A., & Hinchey, M. (2014). An overview of Dynamic Software Product Line architectures and techniques: Observations from research and industry. *Journal of Systems and Software, 91*, 3–23.

DoD. (2017). Systems engineering. In *Defense acquisition guidebook* (Ch. 3). Defense Acquisition University.

Deputy Assistant Secretary of Defense for Developmental Test and Evaluation (DT&E). (2016). *Developmental test and evaluation annual report for FY2015*.

Deputy Secretary of Defense for Systems Engineering. (n.d.). Retrieved from http://www.acq.osd.mil/se/http://www.acq.osd.mil/se/

Elfriede, D., Rashka, J., & Paul, J. (1999). *Automated software testing: Introduction, management, and performance.*

Guertin, N., Schmidt, D. C., & Sweeney, R. (2015). How the Navy is using open systems architecture to revolutionize capability acquisition. In *Proceedings of the 12th Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.

Guertin, N., & VanBenthem, P . (2016). Modularity and open systems architecture applied to the flexible modular warship. *ASNE Journal*.

Hitchins, D. K. (2003). *Advanced systems thinking, engineering, and management*. Artech House.

Hohpe, W. (2004). Enter*prise integration patterns: Designing, building, and deploying messaging solutions*. Addison-Wesley.

Hosain, S. Z. (2016, June 28). Reality check: 50B IoT devices connected by 2020—Beyond the hype and into reality. Retrieved from http://www.rcrwireless.com/20160628/opinion/reality-check-50b-iot-devices-connected-2020-beyond-hype-reality-tag10

How many millions of lines of code does it take? [Graph] (2017, February 8). Retrieved from http://www.visualcapitalist.com/millions-lines-of-code/

Hunt, G., & Allport, C. (2016). Disambiguate data with documentation. Future Airborne Capability Environment Technical Interchange Meeting. The Open Group.

IEEE. (2009). IEEE standard for reduced-pin and enhanced-functionality test access port and boundary-scan architecture (1149.7 Standard).

Jamshidi, M. (2008). *System of systems engineering: Innovations for the twenty-first century.*

Kent, W. (1983). A simple guide to five normal forms in relational database theory. *Communications of the ACM, 26,* 120–125.

Kuhn, Kacker, & Lei. (2013). *Introduction to combinatorial testing*. Chapman and Hall/CRC.

Lethart, R., Porter, A., Schmidt, D., O'Hare, M., Crisp, H., & Laird, B. (2016). *Capability-based technical reference frameworks for open system architecture implementations*. Presented at the SEDC 2016 Conference. Retrieved from http://www.sedcconference.org/capability-based-technical-reference-frameworks-for-open-system-architecture-implementations/

Micouin, P. (2014). *Model based systems engineering: Fundamentals and methods.*

National Defense Authorization Act. (2017). Section 805.

The Open Group. (n.d.). FACE technical standard. Retrieved from http://www.opengroup.org/face

Rahman. (2014). DevOps, continuous delivery (CD) and agile—Is this going to change the world of software engineering? *Continuous engineering for dummi*es. IBM, Wiley.

Richardson. (2016). *Design for maintaining maritime superiority*.

Schmidt, D. C. (2016, July 11). A naval perspective on open-systems architecture [Blog post]. SEI Blog. Retrieved from https://insights.sei.cmu.edu/sei_blog/2016/07/a-naval-perspective-on-open-systems-architecture.html

Tolk, A., & Muguira, J. A. (2003). The levels of conceptual interoperability model (LCIM). In Proceedings of the *IEEE Fall Simulation Interoperability Workshop*. IEEE CS Press.

Decision-Based Metrics for Test and Evaluation Experiments

Dashi Singham—is a Research Assistant Professor of Operations Research at the Naval Postgraduate School, where she researches, teaches, and advises student theses. Dr. Singham's primary areas of focus include simulation modeling, simulation analysis, and applied statistics, with most of her work on developing new methods and metrics for analyzing simulation output. Her areas of application include energy and intelligence systems. She received her PhD in Industrial Engineering and Operations Research from the University of California Berkeley in 2010. [dsingham@nps.edu]

Abstract

We develop a new decision-based metric for determining sample sizes in Test and Evaluation experiments. Traditional confidence intervals for the mean can be used, and we present sequential confidence interval procedures as a way to derive efficient intervals. We discuss decision rules for analyzing the observed output and how to choose confidence interval methods for calibrating these decision rules. The metric presented can help determine if a fast decision on the quality of the system can be made or if many more tests are needed to ensure an accurate estimate of performance relative to a desired standard.

Introduction

Test and Evaluation (T&E) experiments are often conducted with the intent of answering a question about the feasibility of a new system. This system may have properties that are unknown, so rigorous testing is required to ensure the safety and performance of the system before it is adopted. We will use the term "system" to include any object under scrutiny via testing, be it a weapon, computer program, or piece of equipment. This work mainly applies to Developmental T&E where different performance metrics are analyzed individually, though it could also apply to Operational T&E where many varying factors are jointly tested.

In this paper, we will assume that there is a quantifiable non-binary metric for evaluating system performance so that averages and confidence intervals can be easily constructed. The intent of this work is to show how to better use quantifiable metrics to answer research questions or make a decision about the quality of the system in Developmental T&E. We will outline basic metrics for quantifying uncertainty in system output and show how these metrics can be mapped to a decision rule.

The main goal of data analysis is often to estimate the performance of a system using experimental data, sometimes using the sample mean or a confidence interval for the mean as the metric for evaluating the quality of the system. While these metrics are useful, they would be even more useful if they could be mapped directly to a decision. For example,

- If the system mean performance is greater than some value *D*, then we should adopt the system.
- If the system mean performance is greater than some value *D* with probability *x*, then we should adopt the system.
- f a 95% confidence interval for mean performance has a lower bound greater than *D*, then we should adopt the system.

In the examples above, *D* is the decision threshold that is used to determine whether or not to implement the system. It is important to decide beforehand the metrics for success and determine what *D* should be to ensure that the system is selected only if it will satisfy its intended purpose. Waiting to choose *D* until after the system has been tested can lead to bias based on initial test results, and these initial results can be misleading if the system has a high variance.

In this paper, we will describe how confidence interval procedures can be used to design statistical test rules that link the data analysis to the decision threshold. By making a simple change to standard confidence interval procedures, new rules can be developed that incorporate the decision threshold D. These new rules will enable a better determination of whether the system should be implemented and can potentially save testing costs.

Confidence Intervals

A simple way to evaluate the effectiveness of the system is by taking the average of the test results. Let $\bar{x}_n$ be the sample mean of *n* test replications. This average can be compared to the status quo, to the averages of competing systems, or to a decision threshold. However, looking at the average alone does not account for variability and uncertainty in system behavior. Assuming that the system will always perform near the mean when it is implemented could significantly underestimate risk.

Confidence intervals provide a method for assessing the uncertainty in mean results. Let s_n be the estimate of the standard deviation of the data based on *n* samples. Define s_n^2 as the variance estimate based on *n* samples, calculated as

$$s_n^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}{n-1}.$$

The value of s_n^2 estimates the real variance of the system σ^2, which is usually unknown. Let $t_{\alpha,n-1}$ be the *t*-value associated with the *t*-distribution with *n-1* degrees of freedom and tail probability *α/*2. Furthermore, let *η* be the confidence coefficient desired in the resulting confidence interval. This coefficient is usually 90%, 95%, or 99% and *α* is 1-*η*. The Type I error associated with the test is often denoted using *α*. If the data is normally distributed and the variance is estimated, then the confidence interval for mean system performance using *n* samples takes the following form:

$$\left[\bar{x}_n \pm t_{\alpha,n-1}\frac{s_n}{\sqrt{n}}\right]$$

This confidence interval can be compared to the desired system performance *D*, or to confidence intervals for other systems, as will be discussed later. The center point of the confidence interval giving the estimate of mean performance can be compared to *D*, as well as the width of the confidence interval. Formally, we can define the half-width of the confidence interval as

$$\text{half-width} = \left[t_{\alpha,n-1}\frac{s_n}{\sqrt{n}}\right]$$

where narrower half-widths imply less uncertainty in the mean performance of the system. If the assumption of normality in the data is met, repeated collections of confidence intervals from new experiments will result in (1-α)x100% of the intervals including the true mean of the data *μ*, and ideally this value will be around 90%, 95%, or 99%, depending on the choice of *η*.

Confidence intervals help determine the quality of a mean estimate. A narrow confidence interval (small half-width) implies less variability around the estimated system mean and is desirable, while a wide confidence interval makes it more difficult to predict the behavior of the system. When fixing the sample size used for testing ahead of time, there is no control over the half-width.

Let δ be the desired precision in the resulting confidence interval, which is the maximum half-width that is acceptable to the T&E analyst. Smaller values of δ are desirable because narrower confidence intervals provide more precise information on mean performance of the system. Suppose we have an estimate of the standard deviation s, and have some desired upper bound on the precision in our confidence interval δ. Then we can choose the smallest sample size such that

$$n \geq \left(\frac{t_{\alpha,n-1} \cdot s}{\delta}\right)^2, \tag{1}$$

and this sample size will ideally (though not necessarily) yield a confidence interval for μ that has a half-width smaller than δ. This method can be used to estimate a sample size ahead of time that would be needed to produce a small confidence interval. Oftentimes, budgetary constraints are the driving force behind the choice of *n*. A quick comparison between the budgeted number of samples with the ideal choice of *n* using Equation 1 can help determine ahead of time whether the experiment will yield enough precision to get an adequate idea of the true performance.

We note that many methods exist for choosing the sample size for T&E experiments, and guidelines incorporating sampling for different settings are presented in the *Test and Evaluation Management Guide* (2005), the *2010 Integrated Test and Evaluation Handbook* (United States Marine Corps [USMC], 2010), and the *Operational Test and Evaluation Manua*l (USMC, 2013). This work aims to deliver specific sequential sampling techniques that can be used in conjunction with these guidelines to better inform the sample size decision so that appropriate budgetary effects can be considered.

Sequential Sampling

Sequential sampling rules can be an improvement over fixed sample-size testing because they allow for adjustments to the sample size conditional on system performance as it is observed. Thus, after each test is conducted, the cumulative results are aggregated and an estimated confidence interval is computed. The decision to continue testing depends on the confidence interval produced from past samples. Sequential testing avoids the issue associated with Equation 1 where knowledge of the sample variance is required.

For example, if after 30 test runs of a system the confidence interval for the mean is very narrow, it may be unlikely that more runs will produce any additional information or variety in the results. In this case, testing could stop to avoid wasting money on future samples. However, if the confidence interval is quite wide, more tests should be conducted to better understand the uncertainty in the system. Additional tests will narrow the confidence interval to give more precision in the results and will also help better assess the risk in the system. Sequential rules check the confidence interval after each sample, and determine whether testing should continue. In this section, we provide the mathematical notation for understanding sequential confidence interval procedures.

The benefits of sequential sampling can be immediate when applied to a T&E setting. It is cost effective to stop as early as possible and may be wasteful to continue to test after prior tests have established the performance of the system. However, there is a potential for statistical bias associated with sequential sampling, as will be discussed at the end of this section. In this paper, we will not address this statistical bias directly for brevity, but we acknowledge that when sequential sampling stops with only a few test results, there is a high potential for bias in the results. This bias decreases as the number of samples increases.

Another benefit of sequential sampling is that the tester may have no idea ahead of time how many samples are needed to generate a narrow confidence interval for mean performance. If the underlying variance of the system is known, or can be estimated, then a formula such as Equation 1 can be used. But for new systems, the variance is usually not known and must be estimated as data is collected. Thus, it is difficult to know ahead of time how many tests are needed. Sequential sampling removes the need to make this decision and allows the sample size to be variable and adjust to the conditions of the data.

Sequential sampling rules allow for the tester to stop when some specified criterion is reached. This criterion is often a statistical property of the data collected up to that point. The main example we will use is to stop sampling when a confidence interval with a half-width smaller than some precision value can be generated from the data. Instead of a fixed sample size *n*, let n^*be the number of samples collected as the result of a sequential stopping procedure. This value is random, in that it will vary depending on the output values of the test. The values of n^* can be represented using

$$n^* = \text{argmin}_n \, t_{\alpha,n-1} \frac{s_n}{\sqrt{n}} \leq \delta \qquad (2)$$

where n^* is the smallest value of *n* (the first time the criterion is observed) where the half-width of the confidence interval collected with *n* samples is smaller than the desired precision δ. This value of δ is similar to the one used in Equation 1 and represents the allowable uncertainty in the sample mean estimate (the confidence interval). Recall that s_n is calculated as samples are collected, and this will make n^* random and depend on the particular values of the samples observed up to that point. Equation 2 is called an absolute precision rule, because the desired precision in the confidence interval is fixed ahead of time. Another type of rule is relative precision, where the precision can depend on the mean of the data. An example of a relative precision rule is:

$$n^* = \underset{n}{\text{argmin}} \, t_{\alpha,n-1} \frac{s_n}{\sqrt{n}} \leq \delta \bar{x}_n$$

where the required precision of the confidence interval will be smaller for data that have smaller values (when $\bar{x}_n$ is small). Relative precision is useful when the tester does not have any information on what the mean of the data will be, but wants the error in the mean estimate to be within some percentage of the overall performance (for example the half-width should be within 5% of the estimated mean performance value). Note that for both absolute and relative precision rules, the variance must be estimated with each additional sample.

As an example, some specifications for small-arms tests involve absolute precision rules while others involve relative precision. The report TOP 3-2-045 outlines the maximum permissible error of measurement for small arms tests (United States Army Developmental Test Command, 2007). Some metrics require absolute precision in the results (thermograph reading measurement error must be within 0.6 degrees Celsius) while others require relative precision (the viscometer error should be within 0.5% of the full-scale reading). These error values provided are presumed to be two standard deviations over the data, and these values can be used as is or modified to be used as the input δ in a sequential procedure.

Decision-Based Performance

Simulation experiments are often used to help make decisions on whether to implement or modify a system. Because computer models allow for systems that have not yet been constructed to be tested, we can experiment with lower costs than building a physical model. Test and evaluation plans can include simulation system tests, as well as physical tests of real systems.

A common question is "How should we determine what metrics to use in collecting experimental output?" If the system exhibits variable and uncertain behavior, we usually seek to estimate some measure of performance, μ. Confidence intervals are used to measure variability of an estimate. The risk of the confidence interval estimate is measured using its confidence coefficient (η), and the precision is measured using the interval half-width (δ). Estimates of the mean are collected using a sampling rule, and a confidence interval is constructed to help make a decision.

However, the experimental parameters used are often independent of system performance. We ask for the same risk and precision regardless of the data output and even if the output gives mixed results. We would expect a user to want more strict requirements on the precision when the system performance is close to the boundary between deciding to implement or not. Or, a risk-averse individual may want more confidence in a result suggesting that the system be implemented, and may be quicker to decline to implement a system that is unlikely to be better than the status quo.

For example, it may be critical that a system has performance greater than some threshold *D* (recall the examples from the Introduction). If the first set of experiments shows conclusively that $\bar{x}_n > D$ so that it is highly likely that $\mu > D$, then it is not necessary to obtain a narrower confidence interval. However, if $\bar{x}_n$ is close to *D*, the original value of δ might be too wide to differentiate if μ is actually better than *D*. In this case, a smaller value of δ should be used to drive up the number of samples needed. The type of output confidence interval should depend on the potential effect on the final decision to be made. More precise intervals with higher confidence coefficients should be required when the results of system experiments are close to the boundary between implementation or not. Less strict confidence intervals are needed if the system is performing exceptionally well or poorly, in which case the implementation decision is clear.

While we do not know what the true performance of the system is (hence requiring a T&E study), we do know what the decision would be if the true performance were known. We can choose confidence interval parameters based on the type of risk and precision we wish to have for different levels of performance. The confidence coefficient and precision are usually chosen before starting a simulation experiment and are static in that they do not change based on the resulting observations collected. We propose changing the precision parameter depending on the values of the observations collected as the procedure is running. This means we could obtain high-precision results for systems that are close to the decision point *D*, while stopping earlier with less precise results if it becomes clear early in the experiment that the system should not be implemented.

As a way of measuring the effectiveness of sequential confidence interval procedures, confidence interval coverage is often used, where coverage is the proportion of intervals generated by the procedure that cover the true mean μ. Nominal coverage is important for establishing validity of a procedure. However, here we consider the possibility that while a confidence interval may not cover the true system performance mean μ, it still may cover values that would lead to the same decision. There is usually some asymmetry in the type of error the tester will accept. For example, overestimating cost may be better than underestimating cost. But, if the procedure overestimates cost so much that an otherwise profitable system is no longer implemented, then the procedure has failed in two ways: in estimating the true cost and in failing to lead to the correct decision.

Let μ be the unknown true mean performance of the system. The threshold point *D* determines a binary decision for whether or not to implement the system. Suppose higher values correspond to better performance. If $\mu > D$, then we might choose to implement the system, and if $\mu > D$, we might decline to implement the system. The confidence interval can help determine the decision by comparing the values it covers to *D.* For example, if an interval lies completely above *D*, the decision would be to implement the system, while if the interval contained *D* and values below it, then the decision may be to delay or decline implementing the system.

In addition to an interval covering μ, we want to estimate the probability that the procedure results in an interval that leads to a correct decision being made. Consider the following four possibilities for an interval in Figure 1. The interval can either cover (include) μ or not, and it can either lead to the correct decision or not depending on its location relative to *D*.

Cover & Correct	Cover & Incorrect
Fail to cover & Correct	Fail to cover & Incorrect

Figure 1. Four Possible Confidence Interval Situations

Figure 2 illustrates the four situations presented in Figure 1. We assume that the mean performance of the system, μ, is greater than the decision threshold *D*. Thus the "correct" result of the experiment is that the system should be implemented. The top-left plot shows a confidence interval using parentheses that covers the true mean μ, and also lies on the right side of the decision threshold, thus making the correct decision. The top-right figure shows a different confidence interval that also covers the true mean μ. However, it fails to correctly predict that performance is greater than the decision threshold, because the confidence interval includes values on the left and right of *D*. The bottom-left confidence interval fails to cover the true mean μ. However, it is so far to the right that it still correctly estimates performance as greater than *D*. The bottom-right confidence interval not only fails to include μ, but it lies on the wrong side of *D*, so it will incorrectly predict that system performance is worse than *D*.

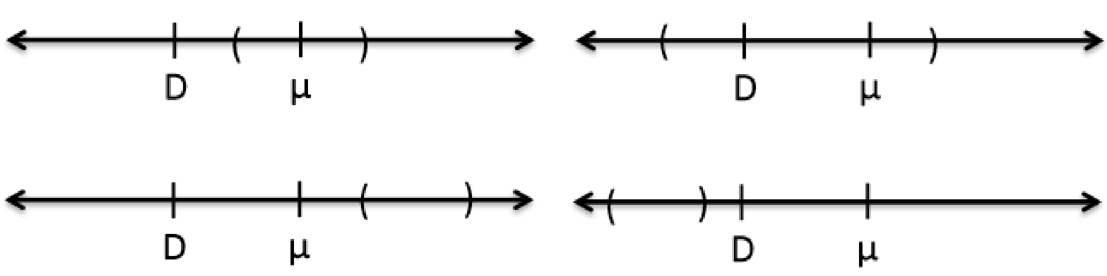


Figure 2. Visual Representation of Confidence Interval Situations

The goal of most confidence interval procedures is to provide adequate coverage of μ so that the procedure produces an interval that includes μ with probability 1-α. However, in the decision context, the correctness of the decision is potentially even more important. Both coverage and correctness are likely correlated, but correctness usually has more of an immediate impact than the effects of confidence interval coverage, which are only realized in the long term.

Implementation

The main result of this analysis is that we should choose δ to be small enough to distinguish μ from *D* in a sequential procedure. A confidence interval that is larger than | μ - *D*| may include μ, but may not be able to distinguish system performance from δ, as seen in the top-right plot of Figure 2. The catch is that we do not know μ at the start of the experiment. However, as samples are collected, $\bar{x}_n$can be used to estimate μ and will be updated with each sample. This value of $\bar{x}_n$ will be the center of each confidence interval. Thus, the sequential stopping rule can be changed to:

$$n^* = \underset{n}{\operatorname{argmin}}\, t_{\alpha,n-1}\frac{s_n}{\sqrt{n}} \leq |\bar{x}_n - D|$$

so that the stopping criterion is such that the experiment will not end until an interval that is small enough to distinguish $\bar{x}_n$ from the decision threshold *D* can be formed.

Making this adjustment would increase the efficiency in standard sequential stopping rules (the absolute and relative precision rules defined above) in a few ways. Sequential stopping rules can be "efficient" because they allow the user to stop as early as possible without wasting effort once a narrow confidence interval has been achieved. However, stopping depends on the choice of δ, which could be arbitrary. What we propose is choosing δ=$|\bar{x}_n - D|$ in an absolute precision rule so that the threshold for the half-width updates and adjusts based on how far away $\bar{x}_n$ is relative to *D*. This way it will be impossible to end with a confidence interval that looks like the top-right plot of Figure 2, because the half-width of the confidence interval will always be smaller than the distance between $\bar{x}_n$ and *D*, so it will never include *D*.

If it turns out the sample mean is close to *D*, then |$\bar{x}_n$-*D*| will be small. This will force the number of samples to increase in order to decrease the confidence interval half-width enough to distinguish whether the system performance is better or worse than D*.* If the sample mean is far away from *D*, then |$\bar{x}_n$-*D*| will be large so it will be easy to meet the stopping criterion after a few samples. Effort will not be wasted when it is clear that μ is on one side or the other of *D*.

The end result is that with this simple change, we can better allocate effort to test systems with a clear idea of the decision threshold. Our decision-making criterion informs the sequential test, and this means we only need to exert the minimum test effort to make a decision. The choice of *D* is very important and should not be made lightly. If $\bar{x}_n$is close to *D*, even if the confidence interval can distinguish system performance, there may still be high levels of risk that require more tests before making a decision about the system.

Of course, standard caveats associated with confidence interval coverage still apply. If too few samples are taken in a sequential procedure, confidence interval coverage can be poor, so the actual confidence could be much lower than the nominal 90% or 95% expected. This is a problem that can be addressed (e.g., in increasing the sample size or changing the expectation in confidence). Chow and Robbins (1965) is the classic reference showing that this bias in coverage decreases to zero as the sample size increases to infinity. However, large sample sizes are often not available in a T&E setting. The other option is to adjust expectations. For example, the tester can run the procedure trying for a 95% confidence interval, while acknowledging that in reality only a 90% interval will be achieved. For more details on calculating and preventing this bias in confidence interval procedures, see Singham and Schruben (2012) and Singham (2014).

Confidence intervals and sampling rules play a major role in determining whether systems meet specified performance thresholds using T&E experiments. For example, in evaluating the performance of body armor in terms of resistance to penetration and deformation, confidence intervals are calculated, and the lower and upper confidence limits are compared to the requirements (Office of the Secretary of Defense, 2010). Specific methods, such as the Clopper-Pearson method, are suggested as a way to calculate confidence interval for probabilities when the output of the experiment is a binary measure of success/failure. We note that other sequential rules may exist for evaluating binary outputs or comparing two hypotheses (Wald, 1973).

While sequential testing may be useful in establishing sampling rules that have the desired precision, adding the decision component *D* would be an easy way of ensuring that the output confidence interval is not only precise but also useful for making the final decision.

Conclusion

Confidence intervals are a useful tool for evaluating T&E data. Sequential confidence interval procedures are a type of sequential testing that determines the sample size by computing a confidence interval after each sample is collected. These procedures potentially allow for a more efficient way of choosing the sample size than fixing it ahead of time. This paper proposes a new type of sequential confidence procedure using a decision threshold that determines whether or not a new system should be implemented based on observed samples. This new metric can potentially be used to either save testing costs or encourage more sampling when system performance is close to the decision threshold. Future work will test the statistical properties of this new metric and simulate the application of decision-based sequential testing using data from past experiments.

References

Chow, Y. S., & Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *The Annals of Mathematical Statistics, 36*(2), 457–462.

Office of the Secretary of Defense. (2010). *Standardization of hard body armor testing* [Memorandum]. Washington, DC: Author.

Singham, D. I. (2014). Selecting stopping rules for confidence interval procedures. ACM *Transactions on Modeling and Computer Simulation, 24*(3).

Singham, D. I., & Schruben, L. W. (2012). Finite-sample performance of absolute precision stopping rules. *INFORMS Journal on Computing, 24*(4), 624–635.

Test and evaluation management guide. (2005, January). Fort Belvoir, VA: Defense Acquisition University Press.

United States Army Developmental Test Command. (2007). *Small arms—Hand and shoulder weapons and machineguns* (TOP 3-2-045). Aberdeen Proving Ground, MD: Author.

United States Marine Corps (USMC). (2010, May). *2010 integrated test and evaluation handbook*.

United States Marine Corps (USMC). (2013, February). *Operational test and evaluation manua*l (3rd ed.).

Wald, A. (1973). *Sequential analysis*. Courier Corporation.

Panel 18. Integrating Program Management and Systems Engineering for Stronger Performance

Thursday, April 27, 2017	
3:30 p.m. – 5:00 p.m.	**Chair: Stephen Townsend,** Director of Network Programs, Project Management Institute ***The Case for Change: The Need for Stronger Engineering Program Performance*** Stephen Townsend, Project Management Institute ***Research of Integrated Engineering Program Management*** Eric Rebentisch, Massachusetts Institute of Technology ***Integration in Practice Case Studies: FA-18 E/F Super Hornet*** Betsy Clark, Software Metrics Jeffrey Morris, National Security Administration

Stephen Townsend—is the Project Management Institute's (PMI's) Director for Global Alliances & Networks, which includes responsibility for alliances with third party organizations, PMI's Registered Education Provider Program, and PMI's Registered Consultant Program. Townsend is also accountable for PMI's program development activities focused on change management, complexity, and implementing organizational project management. Townsend has worked within PMI since 1999 in the areas of member services, chapter/community relations, business/government relations, and PMI's global development activities. Townsend has almost 30 years of experience in non-profit leadership and management. He has also consulted with associations on their global development and community strategies. He is a long-term member of and volunteers within the American Society of Association Executives.

The Case for Change: The Need for Stronger Engineering Program Performance

Stephen Townsend—is the Project Management Institute's (PMI's) Director for Global Alliances & Networks, which includes responsibility for alliances with third party organizations, PMI's Registered Education Provider Program, and PMI's Registered Consultant Program. Townsend is also accountable for PMI's program development activities focused on change management, complexity, and implementing organizational project management. Townsend has worked within PMI since 1999 in the areas of member services, chapter/community relations, business/government relations, and PMI's global development activities. Townsend has almost 30 years of experience in non-profit leadership and management. He has also consulted with associations on their global development and community strategies. He is a long-term member of and volunteers within the American Society of Association Executives.

Eric Rebentisch—is a Research Associate at MIT's Sociotechnical Systems Research Center. He leads the Center's Consortium for Engineering Program Excellence and previously led the Lean Advancement Initiative's Enterprise Product Development research. Dr. Rebentisch's portfolio of research projects include studies of the integration of program management and systems engineering and performance benchmarking of the U.S. shipbuilding industry. [erebenti@mit.edu]

Abstract

Programs to develop and deliver new and enhanced defense systems require strong technical and business management. That requirement means that program managers and chief systems engineers must work closely together as program leadership to enable program team collaboration using aligned tools, practices, and capabilities. While there is plenty of published material focused on enhancing the performance of each individual discipline, very little published matter spotlights how the two disciplines align their efforts and work collaboratively. Extensive research conducted by MIT's Consortium for Engineering Program Management (CEPE), the Project Management Institute (PMI), and the International Council on Systems Engineering (INCOSE) over the last five years has identified opportunities and approaches for improving engineering program management. This paper presents highlights from the research and key factors in integrating systems engineering and program management.

Introduction

> *Taking on large-scale engineering programs is one of the most difficult, risky, and—when done well—rewarding undertaking a government or company can attempt. It not only pushes the envelope of what is possible, but defines a new envelope. It generates capabilities, technologies, products, and systems that are innovative and unique, and generates tremendous societal benefits—from hybrid cars to a trip to the moon, from road networks to GPS navigation, and from carbon-neutral electricity sources to the "smart" city.* (Oehmen, 2012)

So began the text to *The Guide to Lean Enablers for Managing Engineering Programs*, which explored how program managers and systems engineers could impact engineering program performance through collaborative improvement efforts. *The Guide to Lean Enablers* was groundbreaking not just for its application of lean to engineering program management, but also because it has spurred a multi-year conversation and focus on how to build effective inter-disciplinary collaboration capable of solving wicked problems and delivering impactful results.

Engineering programs that incorporate or are reliant upon emerging or evolving technologies are among some of the most challenging to manage. Developing completely novel technologies and then integrating those technologies into systems with other novel technologies requires strong technical and management capabilities. Within the federal government, the U.S. Department of Defense (DoD) has some of the most challenging and expensive engineering programs of any federal agency. In a 2015 report to Congress, the Government Accountability Office (GAO) noted that "The Department of Defense (DoD) has 78 major weapon system programs under way with a total estimated acquisition cost of over $1.4 trillion. These include **some of the most advanced weapons in the world**" (emphasis added). But despite having the most extensive and mature systems engineering and program management capabilities among federal agencies, a 2009 GAO assessment estimated the accumulated cost overrun of the largest 96 engineering programs within the DoD at nearly $300 billion with an average schedule overrun of close to two years (see Figure 1). As the DoD portfolio of state-of-the-art weapon systems are executed through programs that experience extensive cost and schedule overruns, it is clear that the current situation for the DoD is not sustainable.

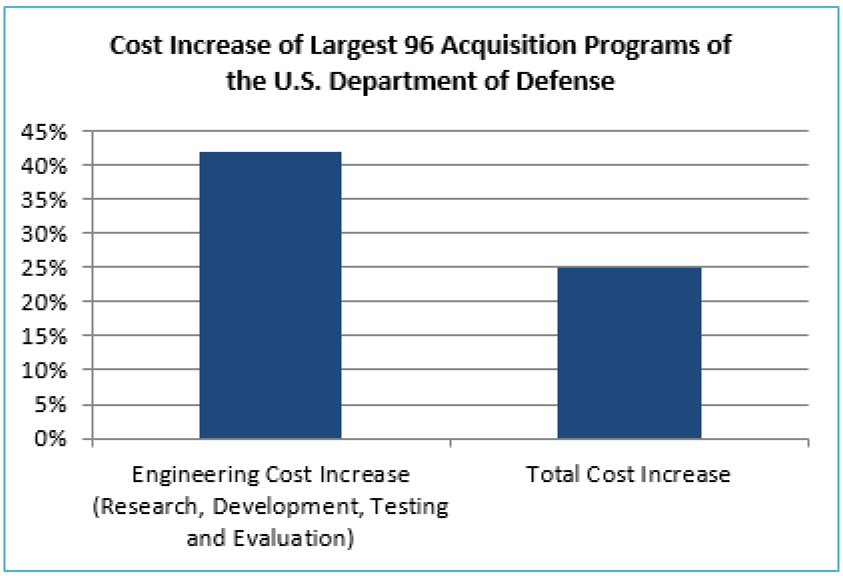


Figure 1. Engineering Programs Are Plagued by Significant Cost Overruns

Demonstrating similar challenges across the federal government, the GAO's High Risk List (GAO, 2017) identifies agencies and program areas that are high risk due to their vulnerabilities to fraud, waste, abuse, and mismanagement or are most in need of broad reform. The GAO's 2017 list includes 34 government operations that are high risk, including a number of agency program areas:

- IT Acquisition Management: "…federal IT investments too frequently fail or incur cost overruns and schedule slippages while contributing little to mission-related outcomes. We have previously testified that the federal government has spent billions of dollars on failed IT investments. These investments often suffered from a lack of disciplined and effective management, such as project planning, requirements definition, and program oversight and governance" (GAO, 2017).
- Department of Homeland Security: The GAO has cited numerous elements of DHS acquisition and program management that need improvement, including

"…tradeoffs stemming from the acquisition affordability reviews; and require components to establish formal, repeatable processes for addressing major acquisition affordability issues" (GAO, 2017).

- National Aeronautics and Space Administration: "NASA manages a portfolio of projects that will always have inherent technical, design, and integration risks because its projects are complex, specialized, and often push the state of the art in space technology. NASA has already taken steps to reduce acquisition risk from both a technical and management standpoint. … However, more needs to be done with respect to anticipating and mitigating risks—especially with regard to large programs, estimating and forecasting costs for its largest projects, and implementing management tools" (GAO, 2017).

Federal authorities recognize the need to transform program performance. The recently enacted Public Law No. 114-264 (Congress, 2016), the Program Management Improvement Accountability Act (PMIAA), outlines specific requirements related to the following:

- Use of standards, policies, and guidelines for program and project management within federal agencies
- A job series for project and program management professionals within the U.S. federal government
- A five-year strategic plan for developing and improving project and program management capabilities
- Establishment of a Program Management Policy Council and portfolio reviews of government programs
- Designation of program management improvement officers
- Adoption and use of best practices in project and program management

Effective transformation efforts like those expected from PMIAA can be better enabled when the key transformation influencers can be identified and leveraged. This paper presents some of the key influencers for better aligning technical and management practices within programs to improve collaboration and drive stronger program performance.

Throughout this paper, the term *program leadership* is used and refers to the technical and management leadership within the engineering program. The majority of research studies upon which this paper is based identified those roles of the program manager as the management leader and the chief systems engineer as the technical leader. Those roles were defined in the following manner:

- *Program manager refers to the job position that has the ultimate authority and accountability* for the overall program.
- *Chief systems engineer* refers to the job position that has ultimate technical authority and accountability for the product or system being developed by the program.

Common Challenges Affecting Engineering Program Performance

In industry and government, there are common challenges that can affect engineering program performance. Research conducted by MIT, PMI, and INCOSE in 2012 explored the application of lean principles to engineering programs in order to eliminate waste and to produce better program performance. Through extensive stakeholder engagement, data collection, and analysis, this research endeavor collected, validated, ranked, and aggregated the most common challenges that affected engineering program performance. The findings consolidated into 10 major challenges (Oehmen, 2012):

- *Insufficient Program Planning*: Program planning may be inaccurate, unable to accommodate uncertainties, or both, which leads to unrealistic expectations and plans.
- *Firefighting*—Reactive Program Execution: The program is executed in a reactive mode toward inside and outside influences, instead of proactively managing and coordinating stakeholders, risks, and issues.
- *Unclear Roles, Responsibilities, and Accountability*: The roles, responsibilities, and accountability of individuals, teams, projects, staff functions, and line functions are not clearly defined in this theme.
- *Mismanagement of Program Culture, Team Competency, and Knowledge*: The expertise and knowledge of individuals, teams, and the organization are insufficient, not transferred properly, or not applied appropriately during the program. It is difficult to establish a productive program culture.
- *Unstable, Unclear, and Incomplete Requirements*: Changing, unclear, and incomplete requirements from customers and other stakeholders seriously affect the efficient and effective execution of the program.
- *Insufficient Alignment and Coordination of the Extended Enterprise*: The complex network of organizations and departments involved in delivering the program value is not aligned to its priorities. This includes the alignment and optimization of strategic priorities and portfolios.
- *Locally Optimized Processes That Are Not Integrated Across the Entire Enterprise*: When processes are only locally optimized, there is a lack of visibility for the value stream, and/or barriers between organizational units to implement a seamless flow. There are insufficient tradeoffs between units to reach an overall optimum.
- *Improper Metrics, Metric Systems, and KPIs*: The metrics and KPIs used during the program do not capture the intended performance attributes, incentivize the wrong behavior, or are lagging instead of predictive.
- *Lack of Proactive Program Risk Management*: Budgetary and time constraints force limited or no risk management activity to be undertaken by the program team. The program team attempts to function without clear off-ramps and mitigation approaches. Ownership of risks is ill-defined.
- *Poor Program Acquisition and Contracting Practices*: Policies and other constraints restrict the program's ability to apply emerging and best practice in complex program acquisition or contracting.

All of the program challenges identified by the lean research are influenced to varying degrees by factors that are external to the program and over which program leadership may have little, if any, control. For example, program leadership may have limited input or influence related to such things as human resource policies or legal and regulatory requirements imposed on programs. Still, program leadership has substantial control over the degree to which some of the above challenges affect the program team, culture, and performance. And while program leadership may not be accountable for such things as advancing best practice across the entire enterprise, leadership can play a proactive role in sharing and facilitating adoption of program management best practices. Table 1 highlights some of the major program challenges from *The Guide to Lean Enablers* over which program leadership may have influence and where strong leadership can mitigate the impact of the associated program challenges.

Table 1. Challenges Internal and External to the Program

Challenges Internal to the Program		Challenges External to the Program	
People Factors	**Process Factors**	**Policy Factors**	**Strategic Factors**
Unclear Roles, Responsibilities, and Accountability	Planning: • Insufficient Program Planning • Firefighting—Reactive Program Execution • Improper Metrics, Metric Systems, and KPIs	Insufficient Program Planning	Insufficient Alignment and Coordination of the Extended Enterprise
Mismanagement of Program Culture, Team Competency, and Knowledge	Lack of Proactive Program Risk Management Unstable, Unclear, and Incomplete Requirements Insufficient Alignment and Coordination of the Extended Enterprise Locally Optimized Processes That Are Not Integrated Across the Entire Enterprise	Unstable, Unclear, and Incomplete Requirements Poor Program Acquisition and Contracting Practices	Unstable, Unclear, and Incomplete Requirements

The remainder of this paper will detail findings from research aimed at exploring the people and process factors over which program leadership can exert influence to help their programs produce stronger results.

Research on Integrated Engineering Program Management

The joint lean research by CEPE, INCOSE, and PMI exploring ways to overcome the 10 engineering program challenges included a second phase of research. That phase collected, validated, and aggregated 43 potential mitigation approaches called "lean enablers" that could be applied to the challenges. That research included specific examples from engineering programs illustrating how the "lean enablers" were applied within actual programs to positively impact program results (Oehmen, 2012). An unstated but underlying presumption in that work was that the program manager and chief systems engineer would lead the application of these "lean enablers" within their programs. However, none of the previous research had explicitly explored how effectively these program leaders collaborated in leading program teams. To address that presumption, a new multi-year, multi-phase research effort was undertaken. The research has culminated in the recent publication of the book *Integrating Program Management and Systems Engineering: Methods, Tools, and Organizational Systems for Improving Performance* (Rebentisch, 2017), which seeks to help program managers, chief systems engineers, and their executive leaders enhance joint effort, joined thinking, and common language.

Integration Research Phase I

Phase I of the integration research was exploratory in nature and was designed to (1) understand how well program managers and chief systems engineers collaborated and (2) detail the degree to which the two disciplines integrate practices from each discipline to effectively manage engineering programs. The research results provided key insights into four areas: roles and authority; use of standards and guidelines; formal alignment of technical and management processes; and the causes and degree to which unproductive tension affected program team performance.

On the point of unproductive tension, differences in the approaches, objectives, and incentives of program managers and chief systems engineers, respectively, to execute their responsibilities can result in tension. This tension can be productive if it forces the different disciplines to share, collaborate, create common understanding, and make tradeoffs in the pursuit of a common set of solutions. Innovation can result from this collision of different perspectives as new ideas are introduced on how to solve a challenging problem. But tension can become unproductive if the parties dig in and cling to their own perspectives about why they are right. The integration study asked individuals to rate the extent to which they had experienced unproductive tension between program managers and chief systems engineers.

Unproductive Tension

As shown in Figure 2, the research uncovered that almost one-third of respondents reported there was unproductive tension between the chief systems engineer and program manager to the point that the tension affected program performance. Slightly more than half (52%) reported minimal unproductive tension that did not substantially affect program performance because the program manager and chief systems engineer were able to work through their problems (Conforto et al., 2013a).

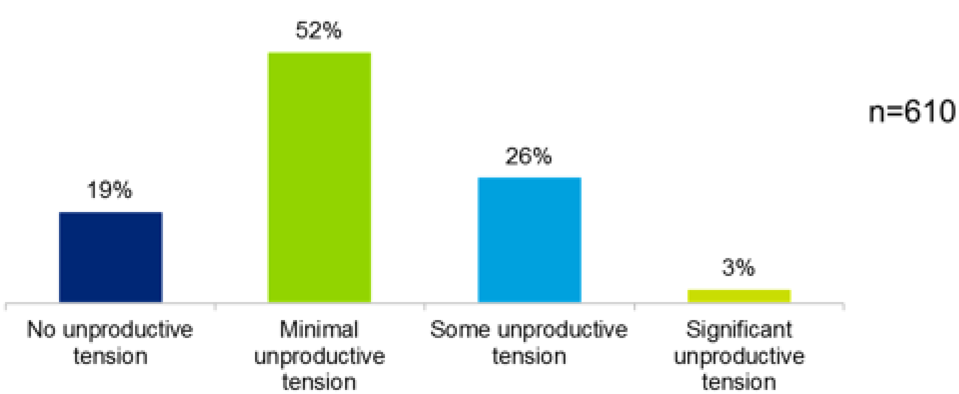


Figure 2. Level of Unproductive Tension

As highlighted in Figure 3, unproductive tension linked back to people and process issues such as unclear roles, lack of planning, and conflicting practices (Conforto et al., 2013a).

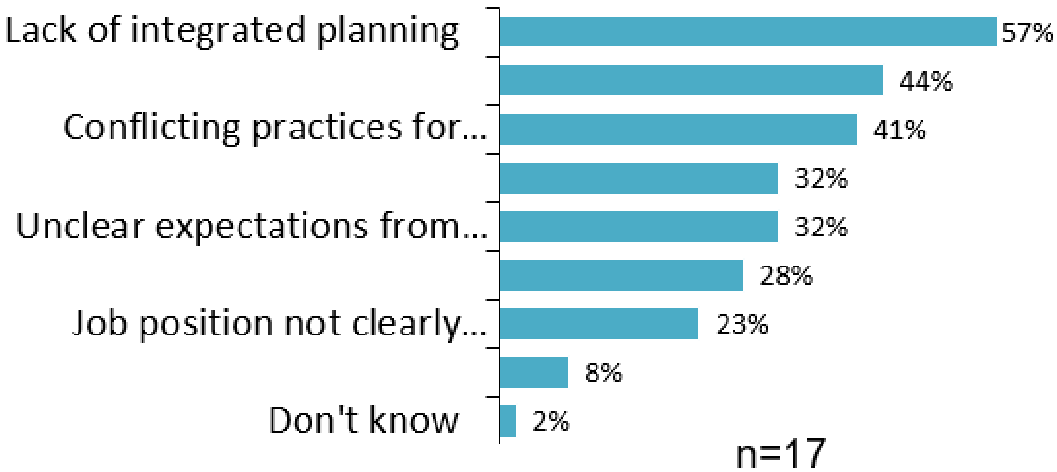


Figure 3. Sources of Unproductive Tension

Roles and Authority

The research found key discrepancies in the degree to which roles and accountabilities were formally defined, whether through position descriptions or in program chartering documentation. The program manager role and authority tended to be more formally defined while both the role and authority of the chief systems engineer were less likely to be formalized, as identified in Figure 4 (Conforto et al., 2013a).

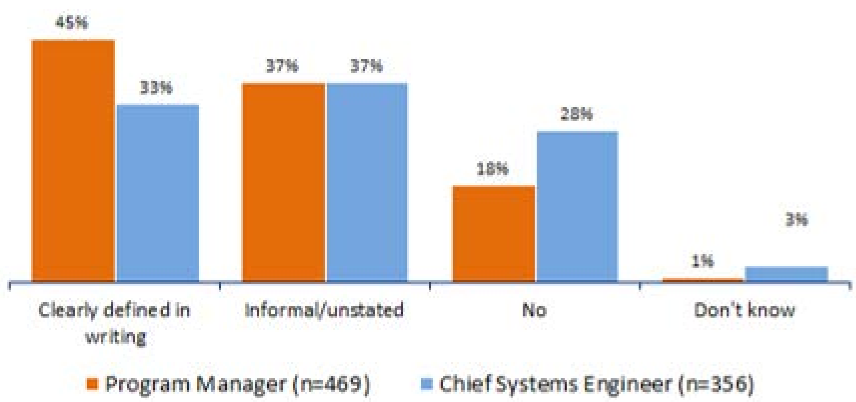


Figure 4. Degree to Which Roles Are Formalized

This discrepancy seemed to influence the chief systems engineers' perception that unproductive tension with the program manager existed because of unclear authority (Conforto et al., 2013a).

The research also indicated that while each role had distinct responsibilities, there were also shared responsibilities in key areas including program/project risk management, external supplier relations, quality management, and lifecycle planning, as shown in Figure 5 (Conforto et al., 2013a). So where role and authority were unclear and where responsibilities overlapped, these factors seemed to contribute to unproductive tension within program leadership.

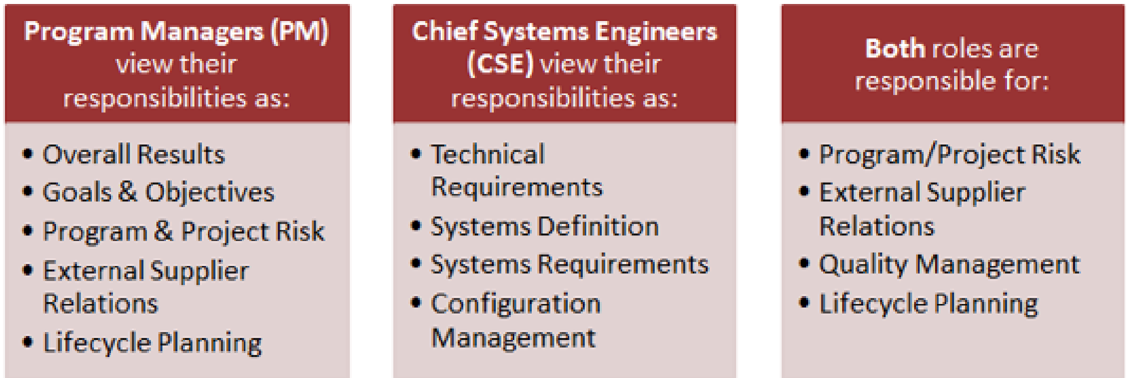


Figure 5. Distinct Roles With Some Overlapping Responsibilities

Integration of Standards

The research found that the majority of chief systems engineers and program managers used domain-centric standards within their programs, as illustrated in Figure 6. It also uncovered that there was not significant use of standards spanning disciplines (Conforto et al., 2013b). So as with unclear roles and authority, the lack of aligned ways of approaching common areas of responsibility was sometimes a contributing factor to unproductive tension within program leadership.

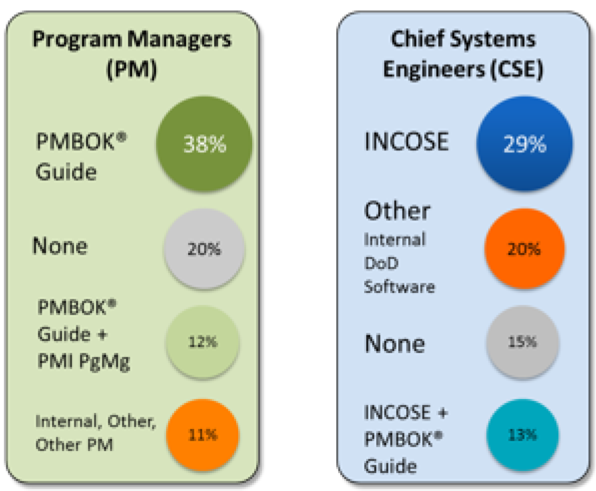


Figure 6. Use of Standards by Each Discipline

How Integration/Alignment Occurs

The research explored how program leadership integrated and aligned practices and standards from the systems engineering and program management domains within their programs. Only 48% reported that program practices were fully or mostly integrated. Where integrated approaches existed, they came about through a mixture of organizational process requirements and of program team members taking the lead to align practices (Conforto et al., 2013a). Both of these results are highlighted in Figure 7.

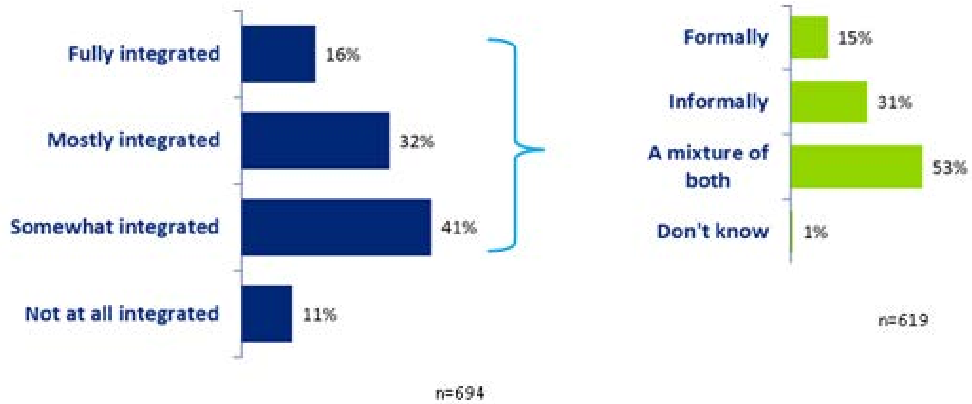


Figure 7. Level and Formality of Integration Efforts

Most respondents (60%) reported that the level of integration was only somewhat effective (Conforto et al., 2013a).

Key Finding From Phase I: Higher Integration Impacts Unproductive Tension

Further analysis was applied to the collected data to determine whether there were statistically significant relationships between integration, unproductive tension, and use of standards. The first analysis explored the use of standards by level of integration and found that the group that used a standard had a higher and statistically significant different level of integration compared with those that did not use any standard. So the research found that aligning inter-disciplinary processes contributes to greater integration between program managers and chief systems engineers (Conforto et al., 2013).

Additional analysis was conducted to explore whether there was a relationship between the level of integration between program managers and chief systems engineers and the formality of the approach to integration. Again, the analysis looked for statistically significant differences by exploring respondents' answers to two questions—one asking about the level of integration and a second asking whether integration occurred formally (e.g., processes transcend the boundaries across the job position) or informally (e.g., people make the integration occur). The analysis found a statistically significant relationship between the formality of the approach for integrating the two disciplines and the level of their integration where higher levels of integration were associated with a greater degree of formality in the approach to integration (Conforto et al., 2013).

Another level of analysis found that integration reduced the level of unproductive tension between the chief systems engineer and the program manager. The analysis found that lower levels of unproductive tension were more likely to exist at higher levels of integration between the chief systems engineer and the program manager. Further, as integration between the program leadership becomes more effective, unproductive tension becomes minimal or non-existent (Conforto et al., 2013).

Integration Research Phases II & III

The next two phases of research sought to understand the key aspects and practices that differentiated organizations with higher integration levels between program managers and chief systems engineers from those with lower integration levels. It also explored the sources and causes of unproductive tension. All of the data collection and analysis aimed to identify how to achieve better program performance by improving integration between program management and systems engineering. Phase II and Phase III research involved in-depth interviews with respondents whose organizations were at each pole—high integration/low unproductive tension and low integration/high unproductive tension. There were nine interviews with respondents who reported no unproductive tension, and there were seven interviews with respondents who reported facing unproductive tension (Rebentisch & Conforto, 2014).

Defining Components of Integration and Unproductive Tension

The interviews explored what the term integration meant to each of the respondents. Using cluster analysis techniques, key themes began to surface that resulted in defining key components of integration as (Rebentisch & Conforto, 2014):

- Having a shared set of objectives defined by the success of the overall effort
- Everyone knowing what those objectives are
- Clarity and understanding around roles and how each role contributes to achieving the objectives
- Respecting the value of the other's role and contribution to achieving the objectives
- Valuing and promoting "collaboration" over "competition"

A similar exercise surfaced key themes associated with defining unproductive tension and its key components which included the following (Rebentisch & Conforto, 2014):

- Failing to communicate and establish a common set of objectives shared by all
- Individuals/groups focused on achieving objectives defined by their own disciplinary identity and/or processes
- Being unable to work together to achieve the globally-superior outcome
- Not valuing the other's role and contributions to achieving the globally-superior outcome

Building Effective Integration

Additional themes surfaced from analysis of the interview data related to effective integration. These themes were stronger in organizations with high levels of integration and low levels of unproductive tension and weaker in those organizations with low levels of integration and high levels of unproductive tension. The emergent themes related to effective integration and their key components clustered into three key factors (Rebentisch & Conforto, 2014):

- Process, Practices and Tools: Encourage continuous improvement and change management through integrated planning and problem solving techniques, use and evaluation of combined practices from each discipline, application of integrated performance measures.
- Organizational Environment: Establish and nurture an organizational environment that builds trust, collaboration, and empowerment to achieve shared goals and objectives with clear roles and accountabilities.
- People Competencies: Utilize engagement, communication and knowledge transfer to promote cross-training and understanding, encourage active listening and recognize the value of multiple competencies and skills.

These factors can be visualized as shown in Figure 8 (Rebentisch & Conforto, 2014):

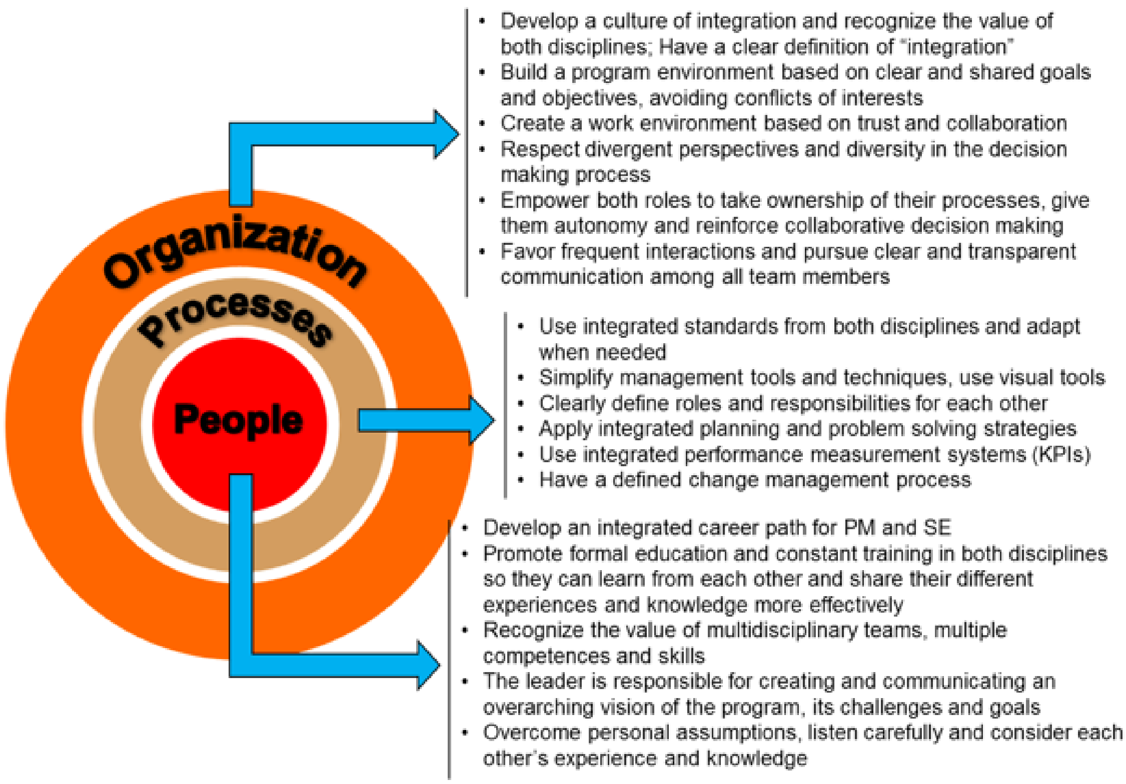


Figure 8. Components of Effective Integration

The positive presence of the components of these themes indicated effective integration.

Conclusions

The complex nature of engineering programs can give rise to significant challenges that can result in cost overruns, schedule delays, poor products, and dissatisfied customers. Program leadership—program managers and chief systems engineers—can greatly influence program performance by collaborating effectively. A key component of strong collaboration involves effective inter-disciplinary integration of people and processes that affect the inner workings of programs. A strong working relationship between the program manager and chief systems engineer enables the type of leadership that can rally a team to overcome hurdles the team might encounter as the program is being executed. The absence of that strong working relationship and the united leadership it provides may exacerbate or amplify the challenges a team encounters while executing a program.

References

Conforto, E., Rossi, M., Rebentisch, E., Oehmen, J., & Pacenza, M. (2013a, June). *Survey report: Improving integration of program management and systems engineering, result of a joint survey by PMI and INCOSE* [White paper]. International Council on Systems Engineering International Symposium, Philadelphia, PA.

Conforto, E., Rossi, M., Rebentisch, E., Oehmen, J., & Pacenza, M. (2013b, June). *Survey report: Improving integration of program management and systems engineering, result* *of a joint survey by PMI and INCOSE*. Presentation at the International Council on Systems Engineering International Symposium, Philadelphia, PA.

GAO. (2015). *Weapon systems acquisitions: Opportunities exist to improve the Department of Defense's portfolio managemen*t (GAO-15-466). Washington, DC: Author.

GAO. (2009). *Assessments of selected weapon programs: Report to congressional committees* (GAO-09-326SP). Washington, DC: Author.

GAO. (2017). *High-risk series: Progress on many high-risk areas, while substantial efforts needed on others* (GAO-17-317). Washington, DC: Author.

Oehmen, J. (2012, May). *The guide to lean enablers for managing engineering programs*. Cambridge, MA: Massachusetts Institute of Technology.

Rebentisch, E. (Ed.). (2017). *Integrating program management and systems engineering: Methods, tools, and organizational systems for improving performance*. Hoboken, NJ: Wiley.

Rebentisch, E., & Conforto, E. (2014, May). Improving integration of program management and systems engineering. Project Management Institute Global Congress EMEA, Session #EM14RES07.

United States Congress. (2016). *Public Law Public Law No: 114-264: The Program Management Improvement Accountability Act (PMIAA).* Retrieved from https://www.congress.gov/114/plaws/publ264/PLAW-114publ264.pdf

Integration in Practice Case Studies: FA-18 E/F Super Hornet

By Elizabeth "Betsy" Clarke—Software Metrics

Abstract

The development program for the FA-18 E/F demonstrated deliberate attention to integrating the program management and systems engineering activities within the program. It also represents one of the few Department of Defense weapon systems programs to finish ahead of schedule, under budget, and with additional functionality beyond original specifications. The program reflected a significant shift from a traditional model of disciplinary stovepipes to a collaborative environment with strongly aligned technical and management leadership. That shift resulted in the following:

- More effective and rapid decision making
- Stronger collaboration and team empowerment
- Clearer alignment of work to product requirements
- Proactive risk management
- Enhanced communication

References

Bailey, E. (1988, April 9). *The F/A-18E/F: An integrated product team (IPT) case study* (IDA NS D-8027). Alexandria, VA: Institute for Defense Analyses.

Bailey, E., Nash, S., & Woolsey, J. (1988, January). *Integrated product and process development case study: Development of the F/A-18E/F* (IDA NS D-2228). Alexandria, VA: Institute for Defense Analyses.

Integration in Practice Case Studies: Electronic Support Upgrade for the Royal Australian Navy's Anzac Class Frigate

By Elizabeth "Betsy" Clark—Software Metrics Inc.

Abstract

This program was an outstanding success, delivering a high priority capability ahead of schedule, within cost, and with minimal defects. It involved the coordination and collaboration of technical staff from seven different companies whose systems and subsystems had to integrate seamlessly. In the early days of the program, the program manager, working closely with the chief engineer, sponsored a series of risk reduction workshops to bring all contractors together to identify key risks and issues and to work together toward their mitigation and resolution. In addition, the program manager and chief engineer fostered an outcome focus on delivering capability to the Navy. This resulted in the following: 1) A high degree of collaboration among all contractors; 2) effective information sharing: Contractors were able to communicate directly without having to channel their communications through third-party bottlenecks; contractors provided each other with computer simulations of their system or subsystem interfaces to allow early integration testing; and 3) rapid and effective decision making in spite of major barriers put in their way.

References

Rebentisch, E. (Ed.). (2017). *Integrating program management and systems engineering: Methods, tools, and organizational systems for improving performance*. Hoboken, NJ: Wiley.

Integration in Practice Case Studies: F-35 Lightning II

By Jeffrey Morris—Lockheed Martin (Ret.)

Abstract

The U.S. Department of Defense's F-35 fifth generation fighter aircraft is the largest development program undertaken by the department and eight partner countries. Effective integration, as experienced on the F-35 Mission Systems software development effort, required experienced leadership, world-class engineers and foundational methods surrounding Earned Value Management (EVM) and Change Management (CM). Organizational and program performance is most effective when program management and engineering functions collaborate during the planning phases. Once the plan is cast, adherence to EVM and CM practices ensure measures are in place capable of predicting deterministic program performance. Post the F-35 Nunn-McCurdy breach, the Mission Systems software re-baseline plan included the following:

- Sensing sessions with individual engineers to assess needs
- Definitive accountability via a revamped EVM process
- Information sharing via a single Change Management system
- Improved workflow via a revised software build and release tool suite
- More effective flight test planning via a new integration process

References

Lockheed Martin. (2017). F-35 fast facts. Retrieved from https://www.f35.com/about/fast-facts

Panel 19. Innovations in Sustainment

Thursday, April 27, 2017	
3:30 p.m. – 5:00 p.m.	**Chair: Lorna Estep,** Director of Resource Integration United States Air Force ***Making Smart Decisions About Supply Chain Security in the Age of Globalization*** Elizabeth McDaniel, Institute for Defense Analyses Michelle Albert, Institute for Defense Analyses Brian Cohen, Institute for Defense Analyses Catherine J. Ortiz, Defined Business Solutions, LLC ***Trends in Performance-Based Services Acquisition*** William Lucyshyn, University of Maryland John Rigilano, University of Maryland ***FAR-Based Crowd Sourcing*** CDR Zachary Staples, USN, Naval Postgraduate School

Lorna Estep—a member of the Senior Executive Service, is the Director of Resource Integration, Deputy Chief of Staff for Logistics, Engineering, and Force Protection, Headquarters U.S. Air Force, Washington, DC. She is responsible for the planning, programming, and budgeting of weapons systems sustainment, equipment, and logistics and installations resource requirements. As part of the Air Force corporate structure, she monitors performance of operations and maintenance, working capital funds and investment programs; participates in program and financial review groups; and advocates for financial adjustments to optimize force readiness. She oversees preparation and defense of these Air Force programs to the Office of the Secretary of Defense, Office of Management and Budget, and Congress.

Estep started her career as a Navy logistics management intern. She has directed the Joint Center for Flexible Computer Integrated Manufacturing, was the first program manager for Rapid Acquisition of Manufactured Parts, and has served as Technical Director of Information Technology Initiatives at the Naval Supply Systems Command. In these positions, she has developed logistics programs for the Department of Defense, implemented one of the first integrated and agile data-driven manufacturing systems, and directed the development of complex technical data systems for the Navy.

As the Director of Joint Logistics Systems Center, Estep had the duties of a commanding officer for a major subordinate command. In addition, she acted as the Logistics Community Manager, an emerging organization to coordinate and implement the revised Defense Department logistics strategy for achieving Joint Vision 2010 through modern information techniques and processes. She has also served as Chief Information Officer for the Naval Sea Systems Command in Arlington, VA.; Executive Director of Headquarters Materiel Systems Group at Wright-Patterson AFB; Deputy Director for Logistics Readiness at the Pentagon; and Executive Director, Air Force Global Logistics Support Center. Prior to her current assignment, she was Deputy Director, Logistics, Air Force Material Command.

Making Smart Decisions About Supply Chain Security in the Age of Globalization

Elizabeth McDaniel—PhD, a Research Staff Member at the Institute for Defense Analyses (IDA) for five years, has focused her efforts on education and awareness of supply chain risk, cyber workforce development, and higher education opportunities for DoD personnel. Prior to IDA, Dr. McDaniel spent her career at institutions of higher education, including National Defense University, as a professor, scholar, and leader. [emcdanie@ida.org]

Michelle Albert—has been a Research Associate at IDA for over three years and serves as a researcher, technical writer, and editor. Her research at IDA has covered a range of topics, including supply chain security and risk management education and awareness. [malbert@ida.org]

Brian Cohen—PhD, has been a Research Staff Member in the Information and Technology Systems Division of IDA for 30 years. Dr. Cohen has performed a range of studies at IDA, with a focus on technology, policy, and business assessments for national security. Recent studies have examined problems with assuring the electronics supply chain for defense systems in the face of increased trends toward offshore sources. [bcohen@ida.org]

Catherine J. Ortiz—is the founder of Defined Business Solutions LLC. She manages the outreach activity for the DoD's Trusted Foundry Program and serves as the Vice Chair of the National Defense Industrial Association's (NDIA's) Manufacturing Division's Supply Chain Network Committee. [cjortiz@definedbusiness.com]

Abstract

Over the last two decades, the Department of Defense (DoD) became increasingly concerned about supply chain security as the supply chain for products became increasingly dependent on commercial and global sources. Supply chains, which are interconnected webs of people, processes, technology, information, and resources around the world, are creating serious asymmetrical threats to our national defense and warfighting capabilities. Hardware- and software-enabled components that traverse these global supply chains afford our adversaries cyberattack vectors that can compromise weapon systems. Educating and enabling the acquisition community to competently assess and make risk decisions in this new area is a particular challenge. Recent education, training, and awareness efforts seek to illuminate a narrow, deeply technical subject such that acquisition professionals can make cost-effective decisions. To that end, this paper presents a new framework for assessing the supply chain risk of particular components while complying with policies and regulations and staying within budget.

Introduction

Electronic components are essential to the Department of Defense's (DoD's) business equipment, communications networks, weapons systems, and supporting platforms. These electronic components follow complex paths, from design, through multiple manufacturing steps, to final test and delivery. The journey that components take through organizations and locations is commonly referred to as a supply chain. Understanding and managing the risks of using components from these generally commercial and global supply chains presents unique challenges for the DoD and its contractors. Attacks by nation-states and other organized groups using the global information and communications technology

(ICT)[1] supply chain can result in service disruption, insertion of malicious functionality, data exfiltration, and intellectual property theft. Supply chain security is imperative for the DoD and other government organizations, as well as many private sector entities. The goal of supply chain security is to reduce a component's or system's susceptibility to supply chain threats[2] and reduce or mitigate the potential impact of any such exploitation.

This paper provides an overview of the risks associated with the global supply chain for the products and systems that power our machines, computer applications, weapons, and vehicles. Threats and vulnerabilities associated with the supply chain create risks that have the potential to affect the performance and security of the components themselves and the systems in which they are integrated. Becoming aware of the risks is the first step; responding appropriately to ensure security, which varies according to one's role in an organization, is the second. Appropriate education, training, and awareness (ETA) and risk mitigation tools must be available for each role. This paper describes the threats and challenges associated with supply chain security and response. The paper capitalizes on the results of an informal survey of current supply chain ETA efforts and offers a new decision-making tool for supply chain security, the Trustworthy Supplier Framework. The evolution and adoption of this new decision-making tool will depend on collaboration across sectors to increase awareness of the problem and to educate and train key personnel across the government, research, and industry communities.

In this paper, *supply chain security* refers to the security and integrity of a component as it travels along its supply chain. *Supply chain risk management (SCRM),* historically considered a logistics-based discipline, focuses on the movement of the component through its supply chain and the threats to this movement, such as earthquakes. The DoD now uses the term *SCRM* in acquisition to refer to the threat of malicious actors who seek to intervene in the supply chain to impair the security of DoD systems and missions. In this paper, the term *SCRM* is reinterpreted and refers to the processes needed to ensure that components are protected against these malicious actors. Much of the policy and guidance discussed in this paper uses *SCRM* in this security context. Both terms are used in this paper, *and supply chain security* and *SCRM* should be considered synonymous and focused on protection against malicious actors.

Supply Chain Security in the Age of Globalization Essentials

Overview of the Benefits and Challenges of Global Supply Chains

The DoD depends on the best and most reliable ICT to build powerful and complex systems. Large systems typically contain thousands of ICT components that are used for

[1] ICT is technology used for gathering, storing, retrieving, and processing information. ICT includes microelectronics, printed circuit boards, computing systems, software, signal processors, mobile devices, satellite communications, and networks. The term ICT reflects the convergence of information technology (IT) and communications; it is not restricted to IT. (IT is defined as any system or equipment "used in the automatic acquisition, storage, analysis, evaluation, manipulation, management, movement, control, display, switching, interchange, transmission, or reception of data" [40 U.S.C. § 11101].)

[2] Throughout this paper we use the term *threats* to refer to nation-state, terrorist, criminal, or other organized actors. Historical supply chain threats such as earthquakes and trucker strikes are also within the scope of threats as used in this paper.

gathering, storing, retrieving, transmitting, and processing information. They include components internal to a system, such as microelectronics, printed circuit boards, computing systems, software, and signal processors, as well as end devices and complete systems, such as mobile devices, satellites, and their networks.

Today, most of the ICT components used in DoD systems and networks are obtained from commercial sources. These commercial products take advantage of global talent, resources, and manufacturing capabilities, resulting in products that typically can be purchased at lower cost than ICT components that are custom-developed for the DoD. Globalization, however, while affording these advantages, creates complex global supply chains that are often opaque and difficult to trace, which creates security challenges. Products traverse national borders and independent companies many times on their way to their point of integration into DoD systems or networks and the end user.

These complex supply chains provide adversaries with a large attack surface within which they can attempt to tamper with, modify, or influence products. The goal of supply chain security is to reduce a component's or system's susceptibility to supply chain attacks and limit any potential impact. Attacks on defense supply chains can occur throughout the DoD system development life cycle,[3] and through multiple entry points. There are entry points for exploitation, manipulation, and counterfeit insertion during component design, manufacturing, testing, transport, delivery, installation, repair, and upgrade.

It is not possible to anticipate or eliminate all vulnerabilities in systems and components, so security risks must be managed and mitigated. And, because everything is connected today, a single exploited ICT component in a DoD system or network can not only affect that system but multiple systems today and in the future. As such, it is imperative to consider the risks associated with each ICT component that is integrated into a system.

Current Guidance for Global Supply Chain Security

Taking action to ensure supply chain security starts with recognizing the threats and vulnerabilities associated with the ICT supply chain, assessing the risks posed by those threats and vulnerabilities, and determining how to manage the assessed risks. When adversaries are successful in their efforts to tamper with ICT components, supply chain attacks create cybersecurity risks that affect a system's confidentiality, integrity, and availability. Supply chain exploitation, a relatively new aspect of cybersecurity, requires the attention of personnel across the system development life cycle, some of whom may be unaware of their critical roles in securing DoD networks and systems.

---

[3] In this document (as well as IDA's ICT SCRM awareness module; see the section titled *IDA's ICT Global Supply Chain Risk Management Awareness* Module for more information), components and systems obtained through simple procurement or as part of the Defense Acquisition Management System (DAMS) are described as having *system development life cycles* from design to disposal. The Joint Capabilities Integration Development System (JCIDS) process is referred to as the *requirements phase.* Some of the DAMS phases have been combined and renamed here for ease of understanding. The *acquisition phase* refers to component and system design, development, testing, production, and deployment; the *operations and sustainment phase* refers to component and system operations and support (including repair or upgrade); and the *disposal phase* refers to the disposal of the component and system.

The DoD recognizes the significance of the threat and is expanding its strategy, articulating new policies, and instituting processes for acquisition, cybersecurity, and risk management to manage ICT global supply chain risk across the system development life cycle. Actions taken in response to supply chain risk will vary depending on the criticality of the system and component and the phase of the life cycle.

Relevant DoD Instructions, Directives, and Regulations

DoD Instruction (DoDI) 5200.39*, Critical Program Information (CPI) Protection Within Research, Development, Test, and Evaluation (RDT&E),*[4] and DoDI 5200.44, *Protection of Mission Critical Functions to Achieve Trusted Systems and Networks (TSN)* (DoD CIO & USD[AT&L], 2012), focus on threats to technology and threats to components, respectively. For protecting CPI, the policy provides guidance to mitigate CPI exploitation; extend operational effectiveness of military systems through the application of appropriate risk management strategies; employ the most effective protection measures, including system assurance and anti-tamper (AT); and document these measures in a Program Protection Plan (PPP). The enclosure on systems engineering in DoDI 5000.02, *Operation of the Defense Acquisition System* (USD[AT&L], 2015), articulates the PPP processes.

The DoD TSN strategy identifies program protection and information assurance implementation as essential to the development of uncompromised weapons and information systems. The strategy strives to integrate robust systems engineering, SCRM, security, counterintelligence, intelligence, information assurance, hardware and software assurance, and information systems security engineering disciplines to manage risks to system integrity and trust. The purpose of DoDI 5200.44 is to minimize the risk that warfighting capability will be impaired due to vulnerabilities in system design or to subversion of mission-critical functions or components. It focuses on mission-critical systems and critical components and suggests risk management processes, tools, and techniques to reduce vulnerabilities, control quality, reduce and mitigate the likelihood of using products containing counterfeit or malicious functions, and increase the traceability of critical components. Systems Security Engineering (SSE), a specialty discipline within systems engineering, supports the development of programs and design-to-specifications that provide life-cycle protection for critical defense resources. The primary vehicle for integrating systems security engineering into systems engineering processes during the system development life cycle is program protection planning. Programs perform criticality analysis to identify their systems' mission-critical functions and components; assess threats, vulnerabilities, risks, and impacts; and select and apply countermeasures and mitigations.

DoDI 8500.01, *Cybersecurity*, instructs the DoD to implement a multi-tiered risk management process that encompasses supply chain risks associated with global sourcing and distribution, weakness or flaws inherent to IT, and vulnerabilities introduced through

---

[4] The DoD defines *CPI* as "elements or components of a research, development, and acquisition (RDA) program that, if compromised, could cause significant degradation in mission effectiveness; shorten the expected combat-effective life of the system; reduce technological advantage; significantly alter program direction; or enable an adversary to defeat, counter, copy, or reverse engineer the technology or capability." CPI includes information about applications, capabilities, and processes and elements or components critical to military system or network mission effectiveness (USD[I] & USD[AT&L], 2015).

faulty design, configuration, or use that will be managed, mitigated, and monitored as appropriate (DoD CIO, 2014a).

DoDI 8510.01, Ri*sk Management Framework (RMF) for DoD Information Technology (IT),* provides the DoD with an integrated, enterprise-wide decision structure for cybersecurity risk management (DoD CIO, 2014b). The framework, captured in National Institute of Standards and Technology (NIST) Special Publication (SP) 800-37, *Guide for Applying the Risk Management Framework to Federal Information Systems: A Security Life Cycle Approach* (NIST, 2010), seeks to improve information security, strengthen risk management processes, and encourage reciprocity among federal agencies. Components, Services, and Agencies are responsible for resourcing RMF implementation. The RMF informs acquisition processes for information technology (IT) and applies to all DoD IT that receives, processes, stores, displays, or transmits DoD information, including information systems (IS); weapons systems; command, control, communications, computers, and intelligence (C4I) systems; sensor systems; and other platform IT (PIT) systems.

The *DoD Program Manager's Guidebook for Integrating the Cybersecurity Risk Management Framework (RMF) Into the System Development Life Cycle* is intended to help program managers integrate cybersecurity into their systems in accordance with the RMF and DoD policy (DoDI 8510.01, 8500.01, 5000.02).

DoDI 5000.02, *Operation of the Defense Acquisition System*, which states the processes and policies for governing the Defense Acquisition System, was updated on February 2, 2017, to include an enclosure on "Cybersecurity in the Defense Acquisition System" (USD[AT&L], 2015). This enclosure discusses a range of cybersecurity risks to DoD systems and networks and assigns program managers responsibility for the cybersecurity of their programs, systems, and networks. The enclosure outlines activities for mitigating cybersecurity risks, including safeguarding program information against a cyberattack, designing and developing systems that can operate in cyber threat environments, and program protection planning. It also discusses specific actions to implement during the materiel life cycle (by phase) and lists resources for performing cybersecurity and related program security activities.

Relevant NIST Publications

In 2017, NIST updated its *Framework for Improving Critical Infrastructure Cybersecurity*, Draft Version 1.1 (referred to as the Cybersecurity Framework), to include cyber supply chain risk management (cyber SCRM). The Cybersecurity Framework provides organizations with a means of identifying and describing their current cybersecurity posture and their target state for cybersecurity, identifying and prioritizing opportunities for moving toward that target state, assessing their progress, and communicating internally and externally about cybersecurity risk. The Framework has three parts: the Framework Core, a set of cybersecurity activities and outcomes; Framework Implementation Tiers, which characterize an organization's cybersecurity risk management practices; and the Framework Profile, which aligns an organization's risks and needs with standards and guidelines (NIST, 2017).

The Cybersecurity Framework identifies communicating cybersecurity requirements to stakeholders as one aspect of cyber SCRM. Another is identifying, assessing, and mitigating products and services that may be compromised or counterfeit, or are vulnerable to malicious tampering. Cyber SCRM activities include determining cybersecurity requirements for suppliers, enacting cybersecurity requirements in contracts, communicating how the requirements will be validated and verified, and determining whether the requirements are met (NIST, 2017).

NIST SP 800-161, *Supply Chain Risk Management Practices for Federal Information Systems and Organizations* (Boyens et al., 2015), adapts the controls listed in SP 800-53, *Security and Privacy Controls for Federal Information Systems* (NIST, 2013), to SCRM. SP 800-161 provides federal agencies with guidance on assessing and implementing risk management processes and controls to manage ICT supply chain risks. It describes ICT SCRM as sitting at the intersection of security, integrity, resilience, and quality. Security refers to information confidentiality, availability, and integrity; integrity refers to the confidence that an ICT product is genuine and will perform as expected; resilience refers to ensuring the ICT supply chain will provide needed products under stress; and quality refers to reducing vulnerabilities in products that may lead to component or system failure or may provide an avenue for exploitation.

SP 800-161 provides ICT SCRM guidance at the organization, mission/business-process, and information-system levels, and it recommends that organizations build their ICT SCRM processes on a foundation of standardized SCRM practices. SP 800-161 covers more than supply chain security and SCRM, but it focuses on information assurance controls. An organization's ICT SCRM plan should focus on managing risk and be able to adapt to threats, respond to internal changes, and adjust to the rapid change inherent to the commercial sector's ICT supply chains. SP 800-161 presents a catalog of 236 controls divided into 17 families: Access Control, Awareness and Training, Audit and Accountability, Security Assessment and Authorization, Configuration Management, Contingency Planning, Incident Response, Maintenance, Media Protection, Planning, Program Management, System and Services Acquisition, Personnel Security, Provenance, Risk Assessment, System and Communication Protection, and System and Information Integrity (NIST, 2010).

NIST SP 800-161 serves as the foundation for the emerging Trustworthy Supplier Framework that is highlighted later in this paper.

Education, Training, and Awareness

The Imperative for Supply Chain Security Education, Training, Awareness, and Guidance

One challenge in addressing security risks associated with global supply chains is to increase awareness of these risks and prepare the system development life-cycle workforce to assess risk and make effective mitigation decisions and actions. Supply chain risk cannot be eliminated—it must be managed. Managing supply chain risk requires personnel who have roles in the life cycle of systems and components, as well as employees of prime contractors and their suppliers, to be aware of supply chain risks, their relevance to their roles, and appropriate responses.

Education, training, and *awareness* are terms that are often used interchangeably and incorrectly. They pertain to different purposes, time horizons, and methods that rely on learning. Efforts to increase *awareness* seek to focus attention on a topic by presenting facts and issues in a manner meant to generate interest and desire for further learning and to shift thinking or level of concern. Awareness efforts include live briefings, online activities, posters and fliers, and articles. *Training*, which is functional and focused on the "how to" aspect, is designed to change behavior by developing specific skills or competencies.

Training outcomes are typically well articulated so that learners know what is expected of them and at what level of proficiency. *Education* is conceptual, strategic, and future-focused. Education seeks to enhance critical and creative thinking and to develop depth and breadth of understanding of principles, concepts, and ideas and their application in novel situations.[5]

In addition to the awareness, training, and education activities that are advancing action in support of supply chain security, a fourth activity, *guidance*, involves experts, researchers, and practitioners sharing and refining standards and best practices. When ETA is used in the remainder of this paper, it includes guidance.

IDA's ICT Global Supply Chain Risk Management Awareness Module

With the sponsorship of the DoD Chief Information Officer's (CIO's) supply chain security effort, IDA developed the ICT Global Supply Chain Risk Management Awareness module in 2014 (McDaniel, Barth, & Albert, 2014). It presents an overview of ICT supply chain exploitation and its potential risk to the DoD, and it summarizes the current array of responses. It is designed to promote awareness of the risks inherent to the ICT global supply chain and to increase understanding of ICT SCRM. The module was designed for DoD personnel and others with responsibility for oversight, risk management, program management, budget, acquisition, system design and development, security, operations, test and evaluation, and system audit. The module leverages available products and processes to allow users to update and modify the content to fit their purpose and audience. IDA developed the content in part through interactions with experts and stakeholders participating in the DoD Trusted Systems and Networks (TSN) Roundtable, a community of representatives from DoD departments and agencies interested in developing and sharing TSN requirements and best practices.

Traditionally, the term *SCRM* primarily refers to logistics, which deals with packaging and delivering products from the manufacturing site to the purchaser. But mastery of logistics does not necessarily equal or include security. ICT SCRM, as defined in IDA's module, refers to the process of identifying critical components and functions, vulnerabilities, and threats to the supply chain, and developing strategies to respond. It focuses on the security and integrity of the products traversing the supply chain, not just how the products traverse the supply chain. The module, which covers SCRM throughout the system development life cycle, is organized around three themes:

1. The New Insider Threat Is Not a Person—It's ICT;
2. Supply Chain Risk Is a Condition to Be Managed, Not a Problem to Be Solved;
3. Take Action to Manage Global Supply Chain Risk.

The module is designed to prompt DoD personnel to care, think, and act in response to the real risks that result from the supply chains of ICT products across the life cycle. The module is available for public release on DVD and comprises an introductory video, a comprehensive narrative report, an accompanying slide set, and a source document repository. The four-minute video is available on the IDA website at https://www.ida.org/SAC/SACResearchDivisions/ITSD/ITSD_Ideas_Home.aspx.

[5] These definitions are adapted from NIST SP 800-16 (April 1998).

The New Insider Threat Is Not a Person—It's ICT

Theme 1 identifies the elements of supply chain security and explains the national security risks associated with global supply chain exploitation. Most of the ICT components used in DoD systems and networks today are obtained from commercial sources and traverse global supply chains that are often opaque and difficult to trace.

As an ICT component traverses its supply chain, it passes from country to country, company to company, and person to person. Each company has its own logistics security standards, and ICT components are often stored or transported in ways that leave them open to tampering and attack. Supply chain attacks can occur throughout the DoD system development life cycle; entry points for exploitation and manipulation include component design, manufacturing, transport, delivery, installation, and repair or upgrade.

Such attacks can result in disruption of service, insertion of malicious functionality, data exfiltration, and theft of intellectual property. The goal of supply chain security is to reduce a component's or system's susceptibility to supply chain threats and the potential impacts of those threats.

Supply Chain Risk Is a Condition to Be Managed, Not a Problem to Be Solved

Theme 2 explains why supply chain risk must be managed, discusses the key concepts of risk management in the context of ICT SCRM, and offers a range of responses to identified risks. Given the generally complex and opaque supply chains of many critical components from commercial sources, it is impossible to remove risk entirely, and attempting to do so can be extremely expensive. Actively managing risk must be considered for every ICT component purchased or integrated into a system. Limits on time and money require the DoD to focus on risks to mission-critical functions, which are functions that, if compromised, could degrade a system's ability to meet its core mission.

If the assessed risk is high, the DoD has four basic responses: treat it, tolerate it, transfer it, or terminate it. Treating the risk means applying countermeasures and mitigations to lessen the consequence of a compromised component or system by incorporating risk management strategies throughout a component or system's life cycle.[6] Transferring, tolerating, or terminating the risk should be considered if it is better to treat the risk at a later time, if there are insufficient resources to treat it now, or if available treatment options do not reduce the risk to an acceptable level. Options to consider in response to an identified risk range from doing nothing, which entails no effort or extra costs up front, to redesigning a system to avoid using a component with unacceptable risk mitigation options, which involves more effort and higher costs.

Take Action to Manage Global Supply Chain Risk

Theme 3 describes the current complex, dynamic, and evolving environment of relevant government and DoD policies, standards, and strategies that guide the management of supply chain risk across the phases of the system development life cycle.

---

[6] According to the Joint Doctrine, *countermeasures* are devices or techniques applied to impair the operational effectiveness of adversary activity. In the context of ICT SCRM, countermeasures prevent adversaries from exploiting supply chain or component vulnerabilities. Mitigations are actions taken to alleviate the risks or effects resulting from vulnerabilities in critical components or systems (*DOD Dictionary of Military and Associated Terms*, 2017).

The DoD has articulated requirements in acquisition policy (DoDI 5000.02, *Operation of the Defense Acquisition System,* and DoDI 5200.44, P*rotection of Mission Critical Functions to Achieve Trusted Systems and Networks [TSN*]), and in cybersecurity policy (DoDI 8500.01, *Cybersecurity*, and DoDI 8510.01, *Risk Management Framework [RMF] for DoD Information Technology [IT*]). In combination, these policies provide guidance on ICT SCRM for DoD personnel.

Although much of the available policy and guidance focuses on the development and production phases of the system development life cycle, most warfighting, intelligence, and business systems, products, and services spend the majority of their existence in the operations and sustainment phase. Risk management is essential during the design and manufacture phases of the system development life cycle because the decisions made have an impact throughout the life cycle, with implications for operations, routine services, maintenance, and planned upgrades or modifications. Decisions made during the acquisition phase can affect the system throughout its life cycle.

Updating Theme 3

The policies, standards, and strategies discussed in Theme 3, current in 2014 when the module was developed, are not comprehensive or well integrated, and do not offer clear guidance to personnel in every role, position, and organization about what to do in response to ICT supply chain risk. Theme 3 discusses the ICT SCRM responsibilities of certain roles and outlines actions for contracting, procurement, operations and sustainment, and disposal, but concrete activities and strategies are not well developed. The updated DoDI 5000.02 provides detailed guidance for program managers to mitigate cybersecurity risks across the acquisition life cycle, including supply chain security (USD[AT&L], 2015). Also, the controls listed in NIST SP 800-161 (Boyens et al., 2015) and the Cybersecurity Framework (NIST, 2017) provide the greater granularity and concreteness needed for action by personnel in government and the private sector.

DoD CIO continues to lead SCRM efforts for the DoD, the broader U.S. Government, and public-private partnerships (both domestic and international) with industry and academia. DoD CIO monitors and leads DoD implementation and improvement of DoDI 5200.44, contributed to the development of NIST SP 800-161, and partnered with NIST to rewrite the Committee on National Security Systems (CNSS) Directive 505, *Supply Chain Risk Management (SCRM)* (CNSS, 2012). DoD CIO, the Department of Homeland Security (DHS), NIST, and the General Services Administration (GSA) sponsor quarterly public-private Software and Supply Chain Assurance (SSCA) meetings to share SCRM best practices and lessons learned. DoD CIO and its FVEY partners (Australia, Canada, New Zealand, and the United Kingdom) sponsor a Supply Chain and Industrial Base Security Tiger Team that addresses SCRM issues from a FVEY and international military and coalition partner perspective. DoD CIO also works with non-profit and standards development organizations to raise awareness about SCRM and improve commercially acceptable global sourcing standards, which affect hardware assurance, software assurance, and assured services.

Status of Current ETA Programs

The authors engaged communities of interest to identify persons involved with ETA activities related to SCRM and supply chain security. Table 1 summarizes the outcomes of interviews conducted in the first quarter of 2017 with representatives of 28 organizations about their education, training, awareness, and guidance activities related to supply chain security. The organizations interviewed include the DoD, other federal agencies, private sector organizations, universities, and one community college. The authors also included guidance activities from organizations that set standards, share information, and/or sell services or products that certify compliance with community best practices. The TSN Roundtable, the Diminishing Manufacturing Shortages and Material Shortages (DMSMS) Working Group, SAE, the National Defense Industry Association (NDIA), and the Aerospace Industries Association (AIA) are examples of such guidance groups.

Interviewees responded to personal invitations from one of the authors and/or an invitation distributed to members of various working groups/communities of practice. One of the authors conducted the interviews and summarized the key points. Some of the organizational activities described in Table 1 reflect a specific focus on supply chain risk, while others have a more general focus on software or cybersecurity.

Table 1. Organizations and Summary Comments

EDUCATION	
National Defense University (NDU), College of Information and Cyberspace	*Strategies for Assuring Cyber Supply Chain Security*, a graduate-level course, focuses on the systems development life cycle, the impact of counterfeits on cyber security, the DoD's trusted foundry, impacts on critical infrastructure, program protection planning and criticality analysis, the Defense Logistics Agency's role, microelectronics, and software assurance. Intended outcomes are the ability to assess organizational risk, develop a plan for increased awareness, and design a SCRM program based on current policies and best practices.
NDU, Eisenhower School of National Security and Resource Strategy	In a DoD acquisition practices- and policies-focused graduate course taught annually, one lesson includes threats to the supply chain (SC), relevant DoD perspectives and policies, and recent research. The intended outcome is increased awareness for better decision-making in the electronics industry. Student learning is assessed through final presentations on selected industries.
University of Detroit Mercy, Computer and Information Systems. Awareness of the importance of SC security is high in the auto industry.	*Secure Acquisition*, a graduate course, is offered annually and is based on best practices from NIST and the International Organization for Standardization (ISO) and research literature. Students are expected to implement a comprehensive, well-defined, organization-wide standards-based acquisition process; customize an appropriate set of acquisition activities for a given organization or project by life cycle phase; and organize, implement, and manage effective acquisition operations for a complex supply chain.
Worcester Polytechnic Institute, MA	*Supply Chain Risk Management* course developed for defense contractors focuses on the risks of counterfeit, tainted parts and products resulting from malware insertion into hardware, firmware, software, and circuit logic. Addresses threats to manufacturing and the integration of systems engineering practices into security engineering, critical infrastructure protection and information assurance, secure manufacturing practices, open software SC assurance, application security, and secure software development.

TRAINING AND AWARENESS	
AXELOS is a British joint venture between the UK government and Capita plc. It owns and nurtures global best practice frameworks and methodologies, including RESILIA, a cyber resilience best practice portfolio of certified training and awareness learning for all staff, for leadership engagement, and that includes a cyber maturity assessment tool.	Organizations and suppliers are being attacked, targeted, and breached, so response and recovery are now as important as detection and protection. An enterprise-wide response from the top must balance risks and opportunities as well as people, processes, and technology. Training is designed to fit personality differences and leverage multiple learning pathways depending on the risks and appetites of organizations and individuals with varied roles along the supply chain.
Black Duck Software helps companies identify and mitigate open source security risks across application portfolios throughout the life cycle and provides actionable, comprehensive lists of security, legal, and operational risks associated with components in use in a company's code base(s).	Increasing awareness about SC security begins with seeing what is in your code and understanding that software security is ephemeral and dependent on continuous monitoring for new vulnerabilities. To mitigate SC risk and leave an acceptable level of residual risk, options for vulnerable components are to rip and replace, patch, punt, and provide compensating controls.
Boeing, an aerospace company and leading manufacturer and exporter of commercial jetliners and defense, space, and security systems, supports airlines and U.S. and allied government customers in more than 150 countries. Boeing maintains a mature, risk-based security program for supply chain software that focuses on evidence and fact-based trust.	To reduce or eliminate security defects in vulnerable software, enterprise-class suppliers certify that their products were tested throughout the system development life cycle to demonstrate they are free of significant defects. Smaller, often third-party single-product suppliers provide evidence from mutually trusted sources that their software is free from significant, known security defects. Purchasing agents receive training on how to implement standards in various software contract types.
Defense Acquisition University (DAU), a corporate university in the DoD, offers courses to military and federal civilian staff and federal contractors. Two faculty experts and curriculum developers developed two supply chain courses that will be ready by the end of 2017. They offer faculty development, perform mission assistance across the DoD, and conduct SC security seminars.	Two online SC-focused courses, 3–4 hours each, offer guidance on how to address SC risks and provide links to guidebooks and polices for managing risk; they do not make recommendations for specific actions because circumstances vary widely. One is an overview of SCRM and SC logistics with a focus on counterfeits and trusted suppliers for understanding of SC risk and how to mitigate that risk from the point of program inception and throughout the life cycle—through assessments done early and often. The second is an introductory-level cybersecurity-focused course that looks at SCRM and counterfeiting of ICT components. Learners who pass the courses earn credits toward annual continuous learning requirements.
The Information Assurance Directorate (IAD) of the National Security Agency (NSA) is tasked with the information assurance mission and providing customers with confidence in cyberspace by identifying and correcting security vulnerabilities before adversaries exploit them without sacrificing the ability to use their systems effectively.	Two years ago, IAD developed a SCRM situational awareness module in information assurance that is now mandatory; a quiz was recently added. The organization also creates guidance for assessing risk in purchasing commercial solutions and a model for incorporating SCRM in capability packages. Guidance for mitigations is derived from security controls in NIST SP 800-53 Rev. 4 and DIACAP 8500.2. A forum exists for information sharing among government, industry, and academia.
Interos Solutions, a Washington, DC–based company, offers technical services, including vendor risk management; cyber and supply chain policy; and program risk management, critical infrastructure security, training, and awareness.	Increased attention to SC risk and cybersecurity, and changes in delivery methods are reported as shifts in traditional SCRM training. Cybersecurity is now integrated into larger enterprise risks. The NIST Cybersecurity Framework underscores that SC risk is part of the C suite's horizontal risk management. Awareness activities are increasingly computer platform-based and designed to push small segments to employees using tools and processes relevant to one's job and focused on specific threats and responses.

Office of the Director of National Intelligence, National Counterintelligence and Security Center	An awareness video is on YouTube. A 15-minute web-based SC fundamentals course is aimed at government employees and private sector partners. In the pipeline is a SCRM blueprint for contractors, with steps to develop a SCRM program and a companion on best practices. Standards for the intelligence community on criticality and threat assessment have been published; information sharing and vulnerability assessment courses are coming.
The Aerospace Corporation, a federally funded research and development center, provides technical guidance and support to government (intelligence community, military, and civil) and commercial customers to assure space mission success.	Classes are offered to customers about SCRM with a mission assurance perspective. They focus on the distillation and integration of government policies, requirements, and standards that are actionable and can be implemented by contractors. They also focus on SCRM recommended best practices to implement, and how to write requirements for contracts with prime contractors to ensure that the requirements will flow down to subcontractors. Additional classes will focus on how to secure the software development environment and the use of open source code for ASIC and FPGA development.
The University of Massachusetts (UMASS), a multi-campus public university that offers undergraduate, graduate, and professional degrees; collaborates with **itSMSolutions**, an online content solutions provider specializing in the delivery of video and instructor-led training solutions.	UMASS is improving its cybersecurity posture by creating training programs that focus on the knowledge, skills, and abilities to operationalize the best practice controls in the NIST Cybersecurity Framework across organizations and supply chains. Based on a NCSF controls factory methodology created by the university CISO, UMASS is partnering with itSMSolutions to offer online materials, video programs, online labs, mentoring services, and testing services designed to teach organizations how to protect their critical information assets and digital services.
U.S. Air Force Materiel Command	A CAC-enabled online course, “Introduction to Protection of Mission Critical Functions to Achieve TSN,” designed for acquisition and sustainment personnel, is ready for launch. It will provide an overview of requirements for mission assurance to support life cycle risk management. The course will increase awareness and knowledge of DoD efforts to field resilient systems through the TSN methodology; promote understanding of TSN terms, policies, and requirements and their importance to success in fielding resilient systems; and instill and maintain a continuing awareness of the ICT supply chain threats and vulnerabilities affecting mission-critical hardware and software.
U.S. Department of Energy (DoE), Office of the Chief Information Officer, SCRM Resource Center	Awareness, linked to cybersecurity awareness efforts, includes two learning modules—one for IT professionals and one for program managers; newsletters and posters; National Cybersecurity Awareness Month; internal postings; a quarterly speaker series; and a mini cross-agency working group of interest. In development is a mandatory program for authorizing officials of IT systems.
U.S. Department of Energy, Office of Electricity Delivery and Energy Reliability in collaboration with the DHS Office of Infrastructure Protection	Focusing on the energy grid and suppliers, DoE collaborates with the Department of Homeland Security to research and develop SC security efforts that can be applied to the nation’s energy infrastructure. For awareness of cyber security and SC risks, webinars on relevant, innovative, and useful approaches to securing the supply chain are broadcast to manufacturers and electrical, oil, and natural gas organizations.
GUIDANCE TO ADVANCE KNOWLEDGE AND PRACTICE	
Electronic Components Industry Association, a membership organization that seeks to promote and improve the business environment for the authorized	SC security is a hot topic. Counterfeits are increasingly sophisticated, as evidenced by the increase in “generic” components to evade trademark infringement, mislabeling to appear to meet updated standards, remanufactured parts

sale of electronic components by manufacturers, their representatives, and distributors.	sold as new, and tainted programmable components. Cross-organizational forums set standards, and share information to keep members up to date.
Exostar, initially a joint venture of aerospace companies, seeks to remove redundancy in the SC process and increase security by protecting interactions with suppliers.	Leveraging identity, order, and access management processes, EXOSTAR's communication channels manage risk, increase cybersecurity, and reduce burden on suppliers and infrastructure. By assessing a suppliers' compliance with NIST 800-171 cybersecurity standards, system owners and prospective suppliers know the suppliers' cyber security posture and readiness.
Hemisphere Cyber Risk Management helps small businesses respond to cyber, legal, and insurance considerations to minimize their exposure to cyber and legal risk.	The focus is cyber risk management services that provide insight to improve cybersecurity investment strategies, and to make business- and cost-justified decisions on cyber risk, thus lowering total cost of ownership.
MITRE, a federally funded research and development center, promotes SC security with clients who are responding to government policies and directives. MITRE supports the DoD in reaching its awareness and compliance goals.	The DoD Risk Management Framework is taking hold at the grassroots level. Awareness has increased significantly, especially with communities with zero tolerance/no failure because of their missions. Intertwined cyber security and SC security require specific responses in implementation of system engineering practices.
NIST, U.S. Department of Commerce, Computer Security Division, conducts research and offers guidance. Cyber SC risk management work began in 2008 for federal departments and agencies with CNCI #11. Published first report on cyber SCRM in 2012 and official guidance for departments and agencies in 2015.	NIST activities range from formal briefings to training, as well as education, guidelines, and standards used by communities of interest and practice. SCRM is at the forefront of awareness in industry. The Cybersecurity Framework draft version 1.1 includes SCRM. Best practices case studies are posted on NIST's webpage, and industry best practices using anonymized lessons learned are in draft. Current research is on metrics, criticality analysis, predictive analytics modeling, and a quantitative approach to analyzing organizational interdependencies and associated risks.
Parasoft helps organizations deliver defect-free software efficiently.	A variety of training services support clients with continuous testing solutions for their organization's workflow to eliminate security risks through detection and prevention.
The Santa Fe Group specializes in thought leadership surrounding third-party risk management across the supply chain by providing expertise to all industry verticals, including critical infrastructure organizations, to mitigate third-party risk.	The membership-driven Shared Assessments Program identifies third-party risks and best practices in cybersecurity, IT, privacy, compliance, information security, and business resiliency controls. The program assists with assessing risk program maturity and provides "trust but verify" techniques and training, including the Certified Third Party Risk Professional program.
Software Engineering Institute, Carnegie Mellon University developed a software evaluation method that has evolved into a widely used multi-level capability and maturity model.	Continuing research on the Resilience Management Model and its higher level of risks in systems of systems, including hardware, software, and services in operational contexts; security engineering risk analysis; and external dependencies (supply chain) are integrated into courses at CMU and the Heinz executive program for CISOs.
Supply Chain Risk Management Consortium, an informal small business team of professionals with varied expertise; supports clients' development of resilient and secure supply chains.	The Consortium offers clients awareness and tools to increase the efficiency of their supply chains while reducing vulnerabilities such as exposure and counterfeits. The Consortium is considering developing a standards-based certification in supply chain risk, resilience, and security that employers will find valuable for their personnel.

The Open Group, a membership organization, developed the Open Trusted Technology Provider™ Standard–Mitigating Maliciously Tainted and Counterfeit Products (O-TTPS).	O-TTPS's best practice requirements for global SC security and COTS ICT product integrity help protect against maliciously tainted and counterfeit products throughout the COTS product life cycle. The O-TTPS Certification Program certifies IT providers (e.g., original equipment manufacturers [OEM], hardware and software component suppliers, value-add resellers and distributors) who conform to the standard. The Forum, a collaboration of government, academia, and the IT industry, develops and maintains the standard, also known as ISO/IEC 20243:2015.
U.S. Department of Homeland Security (DHS), DHS Science and Technology Directorate, Cyber Security Division	In response to the use of open source software and neglected hygiene of software, research is focused on improving, modernizing, and advancing the science of software assurance. Static analysis tools are being modernized to improve the visibility of software for consumers through ongoing collaboration with the NSA Center for Assured Software, NIST, and others.
Underwriters Laboratories standards, risk assessment services, and component certifications contribute to SC and cybersecurity.	Awareness of whether a component/product meets a minimum set of acceptable cybersecurity requirements is driven through publicly accessible cybersecurity certifications of specific manufacturers' products. Many different supply chain stakeholders, including end users, use these certificates. Procurement tools, press releases, white papers, webinars, direct marketing, government outreach, publications, presentations, and customer training are intended to help the full supply chain understand how to use the certifications to establish a baseline of cybersecurity hygiene.

These interviews revealed that some organizations, institutions, and communities increasingly understand supply chain risk and are actively responding with specific guidance and awareness efforts. The activities described in this informal survey focus on increasing awareness rather than changing behavior as an outcome of training. The guidance activities listed in Table 1 describe efforts made by membership and standards organizations and/or working groups that seek to advance the science, technology, and practice of supply chain security.

Supply Chain Security Throughout the Life Cycle

Although supply chain security is a focus of the acquisition phase of the system development life cycle, it is also critical during the operations and sustainment phase. As such, SCRM security must be considered not only during design and manufacturing, for example, but also during the operations and sustainment and disposal phases. Supply chain security cannot simply be evaluated only once in the system's life; it must be continually evaluated throughout the system's operation and sustainment and disposal. Once an acquisition program or system becomes operational, the suppliers, components, delivery processes, and business processes may change. These changes may alter or add supply chain risks. Most fielded systems, products, and services spend the majority of their existence—in some cases up to 80%—in the operations and sustainment phase of the system development life cycle. This means that ICT components are needed long after the authorized components are no longer manufactured. Repairing and replacing these components creates opportunities for counterfeit insertion, as well as other forms of tampering and exploitation. Supply chain security during the operations and sustainment phase of the life cycle calls for monitoring and periodically (or continuously) re-evaluating changes in risk, suppliers, operational environment, and use. Security efforts focused on replacement parts include purchasing parts from suppliers who understand supply chain security and providing information on any changes to the part, the part's operational environment, any vulnerabilities, and software patches to help manage supply chain risk.

The Trustworthy Supplier Framework

Background and Context

The Trustworthy Supplier Framework is a decision-support tool in development under the auspices of the Deputy Assistant Secretary of Defense for Systems Engineering (DASD[SE]) and the DoD Chief Information Officer (CIO) that emerged from the efforts of supply chain security communities of interest. Its purpose is to increase confidence in the security of products purchased from global commercial suppliers by designating qualified suppliers as "trustworthy." The framework could be used in a buyer's evaluation process, either as criterion for defining which suppliers are qualified or as part of the selection criteria. It may also be possible for a third party, either government or in industry, to act as an accreditation organization. Suppliers may find that having their businesses evaluated as being more "trustworthy" makes them more competitive in the DoD market for trusted products and components.

Like every purchaser, the DoD seeks to purchase trusted products and have confidence in the trustworthiness of its suppliers along its supply chain. Nevertheless, what constitutes a trustworthy supplier or product has not been well articulated. The Defense Microelectronics Activity (DMEA) defines a supplier as "trusted" based on the confidence in the supplier's "ability to secure national security systems by assessing the integrity of the people and processes used to design, generate, manufacture and distribute national security critical components" (DMEA, n.d.) such as microelectronics. In this context, a trusted supplier will:

- Provide an assured chain of custody for both classified and unclassified integrated circuits.
- Ensure that there will be no reasonable threats related to disruption in supply.
- Prevent intentional or unintentional modification of or tampering with the integrated circuits.
- Protect the integrated circuits from unauthorized attempts at reverse engineering, exposing functionality, or exposing vulnerabilities (DMEA, n.d.).

Using this definition, the Trustworthy Supplier Framework considers supplier and product trustworthiness to be based on confidence in the people and processes used to design, generate, manufacture, and distribute national security critical components and on evidence that a product is free of vulnerabilities (intentional and unintentional) that could compromise system or mission security.[7]

The evolving framework is a toolbox of vetted commercial standards and practices that suppliers can use to create trustworthy products. Instead of requiring commercial companies to invest in meeting government standards with uncertain financial return, the framework relies instead on familiar commercial standards and practices that will ease the financial burdens of compliance. The result—trusted products—will be comparable, even if the methods are not.

Current guidance speaks to what should be done about supply chain risk, but it does not provide specific steps, processes, or decision-making tools to mitigate risk. NIST SP

[7] This definition was adapted from DMEA's definition of "trusted" and NISTIR 7622.

800-161 guidance includes 236 controls that appear as a menu of options that requires refinement and expert judgment to select those of greatest relevance and importance. Toward that end, the Trustworthy Supplier Framework has the potential to enrich, spark, and supplement ETA efforts by prompting the acquisition workforce to look at their supply chains from a different perspective and use that perspective to enable better decision-making. The framework bridges the gap between relevant policies and standards and actionable controls. The Trustworthy Supplier Framework can also stimulate and enhance education, training, and guidance for the acquisition workforce by engaging learners in their own supply chain security assurance. By applying carefully selected standards, personnel can improve their processes, systems, and products to benefit everyone, thereby sharing responsibility for the increase in quality.

Framework Development

The origins of the Trustworthy Supplier Framework can be traced to discussions at DoD Trustworthy Supplier Working Group meetings and other gatherings of government, industry, and private sector personnel. In 2014, the working group identified a need for a toolbox for DoD acquisition personnel that would include industry standards for low-risk components, Defense Logistics Agency (DLA) supplier and product qualification for moderate-risk components, and DMEA Trusted Supplier requirements for critical components. The working group assigned a core team to define attributes of trustworthiness, develop a strawman framework of trustworthiness attributes, and align existing trustworthiness supply chain qualification approaches with those attributes.

Framework development began with a survey of government and industry standards and practices related to microelectronics SCRM. After a comprehensive review, the core team identified NIST SP 800-161 as the foundation for the Trustworthy Supplier Framework. The team then reviewed each control[8] in SP 800-161 to determine its relevance to determining a supplier's trustworthiness. SP 800-161 was written with a software focus, so the team interpreted control names and descriptions in the context of component acquisitions and hardware to align better with the framework's intent. Of the 236 controls listed in SP 800-161, 78 were found to be relevant.

The team then rewrote the names and descriptions of the relevant controls to fit a component acquisitions and hardware context. The team mapped the rewritten controls to common hardware vulnerabilities and then mapped the rewritten controls to relevant standards, regulations, and practices. The resulting matrix is cross-indexed and detailed, with controls mapping not only to certain standards and practices but also to specific sections of those standards and practices.

The Trustworthy Supplier Framework Approach

The Trustworthy Supplier Framework is a method for developing and applying system security engineering practices and controls to maintain the quality, safety, and security of DoD systems and missions. In the context of the framework, *quality* refers to systems that are available and work when needed, *safety* refers to the assurance that a system failure or error does not cost human lives, and *security* refers to system vulnerabilities and their susceptibility to compromise or exploitation. Each of these is a

[8] Controls are safeguards or countermeasures used to avoid, counteract, or minimize security risks.

system engineering function, and trustworthiness occurs at the intersection of the three. Ensuring quality, safety, and security, however, is a component-level concern as well as a system-level one. Quality, safety, and security concerns at the component level lead to system issues. The Trustworthy Supplier Framework approaches these functions at the component level (individual products and their suppliers) and offers controls to mitigate risks that would affect both the components and the systems they support.

The framework comprises a series of detailed spreadsheets—the rewritten controls and standards matrix—that function as a toolbox that both DoD personnel and commercial suppliers can use to define the needed level of product and supplier trustworthiness, and then select the appropriate controls to achieve that level. The framework also helps the DoD and commercial suppliers determine how best to implement the controls. Generally, controls state what should be done, but not how to do it. Without specific methods or desired outcomes, organizations can unintentionally select procedures that may be expensive, cumbersome, or less effective than desired. The rewritten controls use language familiar to DoD program personnel to facilitate better implementation. These rewritten controls can serve as a rubric that the DoD can use to assess trustworthiness in suppliers and products. For example, in the case in which one measure of trustworthiness might be compliance with framework controls, if company A complies with 28 controls and company B complies with 10, then the DoD can assume company A is more trustworthy.

The framework also helps improve the DoD's decision-making process for acquiring components beyond trustworthiness assessments. Currently, DoD program managers receive multiple sets of requirements for a single product and often do not have a path to satisfy multiple requirements with fewer actions. The rewritten controls in the framework provide a streamlined process, resulting in an integrated compliance path for multiple requirements. The DoD can also use the framework as a roadmap for determining how to engage with industry about current standards and develop new standards. The DoD has developed its own share of standards, but its efforts occurred without a great deal of industrial or commercial participation. Working with industry to develop and refine standards can be mutually beneficial.

Next Steps

Education, Training, and Awareness

Supply chain security begins with awareness of threats and vulnerabilities, and it is followed by informed decision-making that aligns with policies, available tools, and processes. ETA and guidance efforts need to be accompanied by adequate resourcing to implement the DoD's policies.

The education, awareness, and guidance activities discussed in the interviews are advancing supply chain security. As supply chain security tools, approaches, and processes are refined, training programs in supply chain security will become available for personnel in various roles. The Trustworthy Supplier Framework can become one of the drivers of this training. Exercises designed to focus on supply chain security can identify needed training for acquisition, operations, and sustainment personnel.[9]

Evolving the Trustworthy Supplier Framework

The series of spreadsheets that comprise the Trustworthy Supplier Framework is just the first step. The next would be to turn the framework into a tool, perhaps available over a website. This tool could then be tested through pilot programs and studies that measure its effectiveness and cost benefits. Complying with government standards may come with certain high costs. In many cases it would be easier and less expensive for commercial companies to use their own standards and practices or adopt other industry standards. Measuring the cost benefits of doing so would validate the framework and promote its adoption.

The framework can also support the evaluation of the effectiveness of different control implementations, which might extend into a means of evaluating and qualifying or certifying suppliers as trustworthy.

References

40 U.S.C. § 11101, Definitions.

Boyens, J., Paulsen, C., Bartol, N., Moorthy, R., & Shankles, S. (2012, October). *Notional supply chain risk management practices for federal information systems* (National Institute for Standards and Technology [NIST] Interagency Report 7622).

Boyens, J., Paulsen, C., Moorthy, R., & Bartol, N. (2015, April). *Supply chain risk management practices for federal information systems and organizations* (NIST Special Publication [SP] 800-161).

Committee on National Security Systems (CNSS). (2012, March 7). *Supply chain risk management (SCRM)* (CNSS Directive 505).

Defense Microelectronics Activity (DMEA). (n.d.). Trusted accreditation. Retrieved from http://www.dmea.osd.mil/trustedic.html

Defense Science Board. (2017, February). *Report of the Defense Science Board Task Force on cyber supply chain.*

Department of Defense Chief Information Officer (DoD CIO). (2014a, March 14). *Cybersecurity* (DoD Instruction [DoDI] 8500.01).

DoD CIO. (2014b, March 12). *Risk management framework (RMF) for DoD information technology (IT)* (DoDI 8510.01).

DoD CIO & Under Secretary of Defense for Acquisition, Technology, and Logistics (USD[AT&L]). (2012, November 5). *Protection of mission critical functions to achieve trusted systems and networks (TSN)* (DoDI 5200.44).

[9] The Defense Science Board Task Force on Cyber Supply Chain released a report in February 2017 that recommends that military service chiefs conduct at least one Cyber Awakening exercise per year and use the results to drive and update training. Cyber Awakening exercises identify vulnerabilities and monitoring activities and promote vulnerability-related information sharing (Defense Science Board, 2017).

DOD Dictionary of Military and Associated Terms. (2017, February). Retrieved from the Joint Electronic Library: http://www.dtic.mil/doctrine/new_pubs/dictionary.pdf

McDaniel, E., Barth, T., & Albert, M. (2014). *Managing information communication technology global supply chain risk—Awareness module*. Alexandria, VA: Institute for Defense Analyses.

National Institute for Standards and Technology (NIST). (2010, February). *Guide for applying the risk management framework to federal information systems: A security life cycle approach* (NIST SP 800-37 Revision 1; updated June 5, 2014).

NIST. (2013, April). *Security and privacy controls for federal information systems* (NIST SP 800-53 Revision 4).

NIST. (2017, January 10). *Framework for improving critical infrastructure cybersecurity*. Draft Version 1.1.

Ross, R., McEvilley, M., & Oren, J. C. (2016, November). *Systems security engineering: Considerations for a multidisciplinary approach in the engineering of trustworthy secure systems* (NIST SP 800-160).

Under Secretary of Defense for Acquisition, Technology, and Logistics (USD[AT&L]). (2015, January 7). *Operation of the defense acquisition system* (DoDI 5000.02). Washington, DC: Author.

Under Secretary of Defense for Intelligence (USD[I]) & USD(AT&L). (2015, May 28). *Critical program information (CPI) protection within research, development, test, and evaluation (RDT&E)* (DoDI 5200.39).

Trends in Performance-Based Services Acquisition[1]

William Lucyshyn—is the Director of Research and a Senior Research Scholar at the Center for Public Policy and Private Enterprise in the School of Public Policy, University of Maryland. In this position, he directs research on critical policy issues related to the increasingly complex problems associated with improving public sector management and operations, and on how government works with private enterprise. Current projects include modernizing government supply chain management, identifying government sourcing and acquisition best practices, and analyzing DoD business modernization and transformation. Previously, Lucyshyn served as a Program Manager and the Principal Technical Advisor to the Director of the Defense Advanced Research Projects Agency (DARPA) on the identification, selection, research, development, and prototype production of advanced technology projects.

Prior to joining DARPA, Lucyshyn completed a 25-year career in the U.S. Air Force. He received his bachelor's degree in Engineering Science from the City University of New York and earned his master's degree in Nuclear Engineering from the Air Force Institute of Technology. He has authored numerous reports, book chapters, and journal articles. [lucyshyn@umd.edu]

John Rigilano–is a Faculty Research Associate at the Center for Public Policy and Private Enterprise (CPPPE). He received his Master of Public Policy degree from the University of Maryland's School of Public Policy in 2011, where he also served as a Graduate Research Associate at CPPPE. During his time at CPPPE, he has coauthored and contributed to numerous reports. In 2004, prior to his time at UMD, Rigilano received his Bachelor of Arts degree in Anthropology from Penn State University. [jprig@umd.edu]

Abstract

Performance-based services acquisition (PBSA) is a proven strategy that reduces costs and improves the quality of service. Rather than specify inputs or service requirements, the DoD stipulates a level of performance that the contractor is then obligated to meet or exceed. When used appropriately, this strategy aligns the objective of the contractor with that of the government customer and can increase the effectiveness and efficiency of the services provided.

Recognizing the benefits, the DoD has sought to increase the appropriate use of PBSA. In 2000, the Under Secretary of Defense for Acquisition and Technology directed that 50% of service acquisitions be performance-based (measured in dollars and actions) by the year 2005 (Gansler, 2000). Through analysis of data from the Federal Procurement Data System (FPDS), this report presents trends in PBSA over the last 15 years

Introduction

Performance-based services acquisition (PBSA) is a proven strategy that reduces costs and improves the quality of service. Rather than specify inputs or service requirements, the customer stipulates a level of performance that the contractor is then obligated to meet or exceed. The contractor has the freedom to meet the objective using its resources and personnel to improve processes and effectiveness. This strategy aims to align the objective and incentives of the contractor with those of the customer. When properly structured, these contracts incentivize providers to improve their efficiency.

[1] *This is a preliminary report. The final version will be released in June 2017.

In 2000, then-Under Secretary of Defense for Acquisition, Technology, and Logistics Jacques Gansler issued new guidance: “It is the policy of the Department of Defense that, in order to maximize performance, innovation and competition, often at a savings, perfomance-based strategies for the acquisition of services are to be used wherever possible” (Gansler, 2000). He went on to state that “in order to ensure that the DoD continually realizes these savings and performance gains, I establish, at a minimum, that 50% of service acquisitions, measured both in dollars and actions, are to be performance-based by the year 2005.”

In the ensuing decade, PBSA dollars rose, both as a share of total contract dollars and in absolute terms. This paper will examine trends in PBSA in the absence of any new directives. To what extent has the DoD and its constituent organizations continued to rely on PBSA? Is current use (extent and implementation) of PBSA appropriate? How can PBSA use be improved?

DoD Contracts Spending

The DoD contracts for a large variety of services, ranging from building maintenance to weapons design, healthcare, education, transportation, and food services. In fact, over half of federal acquisition dollars are spent by the DoD. The FAR (2016) defines a contract for services as an agreement “that directly engages the time and effort of a contractor whose primary purpose is to perform an identifiable task rather than to furnish an end item of supply” (DoD, 2009). Figure 1 shows the industries in which DoD services acquisition is concentrated.

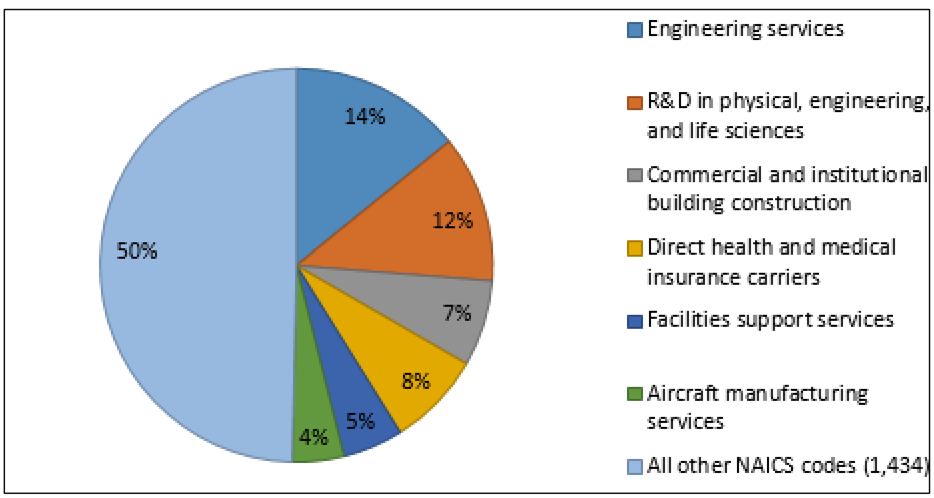


Figure 1. FY2015 DoD Service Contracts Spending by NAICS Code
(Action Obligations [$]; Analysis of FPDS Data)

In fiscal year (FY) 2015, the DoD obligated $275 billion. Of this amount, $130 billion (47%) was spent on non-service contracts (supplies); $145 billion (53%) was spent on services, a figure that includes contracted R&D (see Figure 2). In the 1980s, DoD spending on service contracts averaged only 39% (General Services Administration, 2009). Given the current and projected magnitude of spending on contracted services, improving the efficiency of their acquisition is of utmost importance, especially given continued concern over the DoD’s current acquisition policies. Critics point to growing numbers of “undefinitized contracts,” large numbers of cost-based contracts, the lack of adequate metrics, a general lack of coordination with regard to the procurement of services, and a lack of confidence that the DoD is optimizing the value received from these contracts (House of Representatives Committee on Armed Services, 2009).

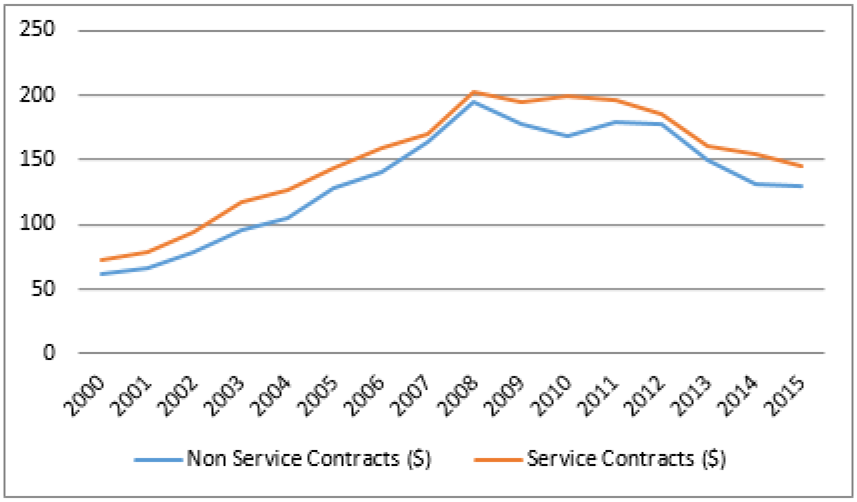


Figure 2. DoD Contracts
(Action Obligations in Billions of Dollars; Analysis of FPDS Data)

PBSA Defined

The broad application of PBSA in the federal government, state and local government, and the private sector has produced many definitions of performance-based contracting. The Federal Acquisition Regulation (FAR) defines PBSA as “an acquisition structured around the results to be achieved as opposed to the manner by which the work is to be performed” (FAR 2.1-13). The Department of Defense guidebook says PBSA “involves acquisition strategies, methods, and techniques that describe and communicate measurable outcomes rather than direct performance processes” (DoD, 2000). The definition used by the National Institute of Governmental Purchasing (NIGP) adds an important distinction: compensation. Performance-based contracting “is a results-oriented contracting method that focuses on the outputs, quality, or outcomes that may tie at least a portion of a contractor’s payment, contract extensions, or contract renewals to the achievement of specific, measurable performance standards and requirements” (NIGP, 2012). Figure 3 shows the proportion of PBSA within the context of overall DoD contracts spending.

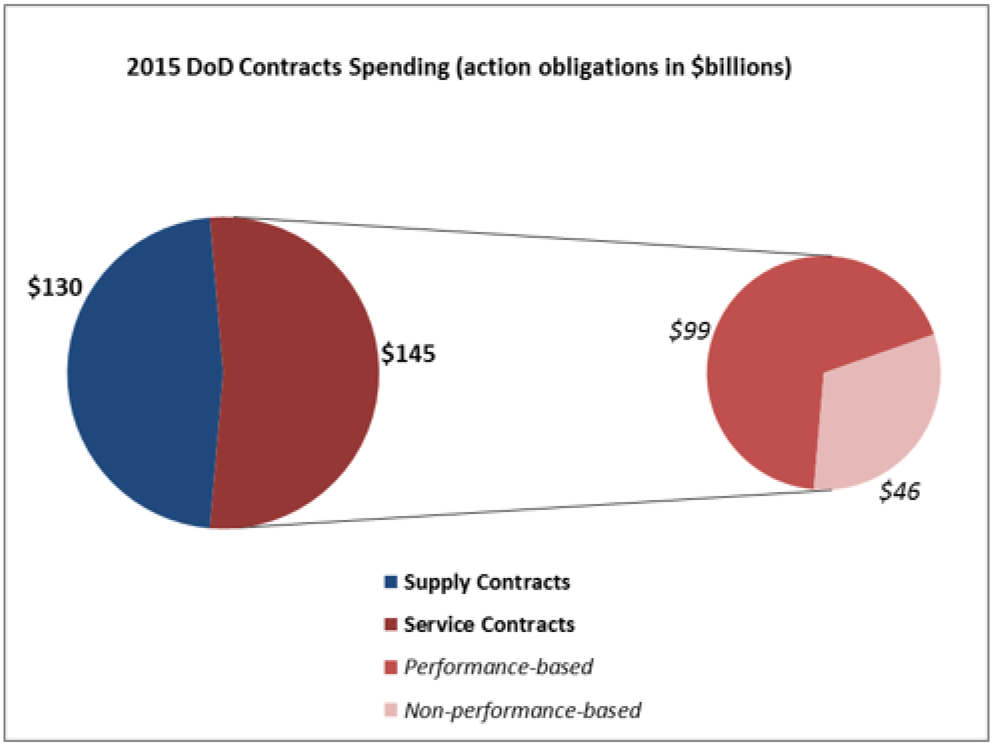


Figure 3. 2015 DoD Contracts Spending
(Action Obligations in Billions of Dollars; Analysis of FPDS Data)

Theoretical Basis

PBSA contracts specify a desired result, without specifying how the work must be performed. This method of contracting diverges from traditional contracting approaches (called compliance contracting or regulatory contracting), which do include narrow specifications on how the result is delivered. Such restrictions minimize the incentive to innovate more efficient processes or products, since the contractor would not profit from such innovations.

PBSA, instead, permits greater flexibility. Contractors are free to pursue efficiencies and innovations that will reduce the cost of meeting the contract's requirements.

The goal of a PBSA arrangement is to align the incentives of the suppliers with the purchaser, so that what benefits one also benefits the other. Consider the interest of a mechanic. Instead of paying a mechanic to repair a car any time it has a mechanical problem, a driver pays the mechanic a fixed sum annually to keep the car operational. This rearrangement produces results by "changing the rules of cooperation so that the self-interested rational choices the agent is likely to make fulfill the outcomes that the principal desires" (Taylor & Shaver, 2010). Such an arrangement of incentives discourages suppliers from performing behaviors which are beneficial to themselves but diminish the quality or availability of the service delivered.

There are other benefits to PBSA as well. For one, it may offer a resolution to the "historic disconnect between the motivation for governments to contract and how they actually go about contracting" (Martin, 2016). This disconnect arises when governments contract for services with the expectation that the superior efficiency of private firms will deliver those services more cheaply and reliably. However, by employing the rigid design specifications (which tell the contractor how to do the work) common to non-PBCs (non-performance-based contracts), governments hinder the contractor's ability to innovate and thereby minimize or negate the private firm's initial advantage, since the strictness of such contracts may disallow a contractor from exploiting whatever innovations and efficiencies they may develop. Such a contractual arrangement is, in part, self-defeating. By instead embracing performance specifications, PBCs (performance-based contracts) allows contractors to profit from reduced costs or innovation. Furthermore, reducing the focus on specifications benefit also reduces public expenditure by minimizing the need for oversight by government personnel.

In the private sector, certain industries have embraced the use of PBCs. Commercial airlines, for instance, were among the first to do so. PBSA in this industry took the form of "power-by-the-hour" contracts, in which aircraft engines and maintenance are provided for a fixed sum per flight hour that the engine is in use, rather than as a fee for the service of engine maintenance. Previously, the engine manufacturers had less incentive to perform preventive maintenance, since they stood to gain from more lucrative maintenance in the future. It is important to stress that the incentives involved can be powerful. For example, Dennis and Kambil (2003) found that in 2003, General Motors' profit rate on after-sales revenue was much higher than that earned through the sale of its cars. In contrast, under fixed sum per flight hour schemes, manufacturers only receive payment when the engine is in use, thereby rewarding reliability. This strategy ensures the engine is more available and at lower cost.

The principles of PBSA have led to reforms in the health care industry as well, under the guise of pay for performance. Pay for performance introduces financial incentives to medical personnel to achieve more optimal patient outcomes rather than be compensated strictly for services performed. The similarities are evident. Furthermore, the clear links between private industry health care and public health care shows that PBSA concepts work in both sectors and between them. The Center for Medicare and Medicaid Services sponsors a Value-Based Purchasing system, intended to pay "for inpatient acute care services based on the quality of care, not just quantity of the services they provide" (Center for Medicare and Medicaid Services, 2015).

Benefits

The benefits of PBSA have been espoused by numerous government and private-sector organizations. We have aggregated and summarized these benefits as follows.

Improved Performance

PBSA helps align the objectives of the contractor with those of the government. Contractors, tasked with achieving outcomes as opposed to fulfilling tasks, (1) have the freedom to implement the strategy that would provide best value to the customer, (2) can update their methods without the need to change contractual obligations, and (3) have the incentive to achieve their best performance. These conditions foster the best effort and innovation on the part of the contractor, maximize the potential for the government to receive optimal contractor performance, and result in a "win-win" for both the government and the contractor.

Lower Cost

Top commercial firms have used PBCs to reduce costs of services even as they raise performance. The federal government, unlike the private sector in its budgetary processes, is not focused on profits; rather, it is focused on transparency; minimizing fraud, waste, and abuse; holding public servants accountable; and costs. The federal government thus often retains more cost-inefficient practices and processes, and will significantly benefit from PBSA's cost savings.

Increased Innovation

PBSA encourages innovation by granting firms flexibility to determine the processes they use to perform the required function. Since they are incentivized throughout the contract to meet the required metrics while minimizing the cost, competitive firms will continuously innovate to improve their processes while reducing costs.

Greater Use of Commercial Services

As noted in a memo issued by the Office of the Deputy Under Secretary of Defense for Acquisition Reform, "the vast majority of service requirements are commercial in nature" (Gansler, 2000). Although government policy explicitly embraces greater use of commercial off-the-shelf technologies and commercial standards, the DoD has been slow to fully implement these policies. By focusing on performance over process, PBSA helps to reduce barriers to entry for commercial firms.

More Effective Oversight

Traditionally, the DoD has spent a large amount of resources verifying that contractors comply with the detailed processes and procedures the government specifies in its contracts—regardless of whether such compliance produces better outcomes. For over a decade and a half, the DoD has been committed to reforms that "ensure that oversight and review of contract management add value to the process and are minimally intrusive" (DoD, 1995). With the performance-based contract structure, the government can reduce the cost and increase the effectiveness of its oversight by tracking appropriately selected performance metrics to monitor contractor performance.

Greater Contractor–Government Cooperation

DoD services are provided through an ever-widening network of contractors. Through several attributes previously described, PBSA encourages a greater contractor–government partnership that is more collaborative and less adversarial than traditional contracting, which implies that companies cannot be trusted to provide a service without being told how to do it. PBSA, on the other hand, is predicated on trust and accountability. Private companies are given more flexibility to find cost-effective solutions, and also agree to meet the required performance metrics, which are often used to determine incentives.

Greater Agility

Contracting for services affords a greater surge staffing capability, giving the DoD a cost-efficient way to augment capabilities during times of increased demand. On the other hand, during times of decreased demand, the DoD can quickly save operating costs by reducing its reliance on service contractors, something not possible with full-time government employees. Moreover, when contracting for services, there is no long-tail cost: the DoD does not have any financial obligation to contractors once the service is delivered or no longer required. Service contracting can also provide the DoD with quick access to required expertise; by contrast, the time required for the DoD to advertise a job position, review applications, perform job interviews, and make job offers is often considerably longer.

Drawbacks

Drawbacks, both real and perceived, have also emerged.

Perception That the Government Has Less Control

Critics of PBSA argue that the government, by not issuing explicit specifications, will have less control, and as a result, could receive less satisfactory performance. This has been shown not to be the case, as the government must identify its critical desired outcomes and then identify the appropriate performance metrics necessary to incentivize the contractor. In many ways this is a superior way of managing outcomes than the traditional method, which has proven to be highly inefficient.

Questionable Applicability

Several critics of PBSA argue that this strategy can only be used for certain types of services. Most of these critics argue that PBSA is best used for contracts that include "many common, routine, and relatively simple services" (Edwards & Nash, 2006). PBSA should not be used when objectives "are too long-term and complex to permit complete specification of results and competitive pricing at the outset of contracting" (Edwards & Nash, 2006).

Ineffective Metrics

Appropriately chosen metrics (1) direct contractor efforts and (2) provide effective oversight. Although concern for appropriate metrics is valid for all DoD contracts, ineffective metrics particularly undermine PBSA contracts because they form the basis of evaluating contractor performance. Metrics and corresponding incentives help align the interests of the contractor with the government. If the two are not aligned because metrics misdirect contractors towards unimportant services, then such contracts will be implemented with suboptimal results. Additionally, the government's oversight must rely on accurate, independently verified data. In many cases, however, the contractors usually furnish the government with this data, presenting a potential conflict of interest. For the incentives to be effective, the government must have reliable data that it can use to provide oversight of a contractor's performance.

PBSA Timeline

In 1991, the Office of Federal Procurement Policy (within the Office of Management and Budget) issued Policy Letter 91-2, which ushered in the formal adoption of PBSA by government. The letter declared "It is the policy of the Federal Government that (1) agencies use performance-based contracting methods to the maximum extent practicable when acquiring services and (2) agencies carefully select acquisition and contract administration strategies, methods, and techniques that best accommodate the requirements" (OFPP Letter 91-2, 1991). Subsequent federal legislation like the Government Performance and Results Act (GPRA) of 1993, the Federal Acquisition Streamlining Act, and the Federal Acquisition Reform Act of 1995 formalized this commitment.

The Federal Acquisition Regulation was not amended to incorporate PBSA policies contained in OFPP's Policy Letter 91-2 until 1997 (GAO, 2008). FAR Part 37 provides the DoD with the policy and procedures that are specific to the acquisition and management of contracted services. This Part also identifies performance-based acquisition as the DoD's "preferred method for acquiring services… [which should be used] to the maximum extent practicable," except in certain circumstances. FAR Part 37 also states that the DoD should facilitate greater use of PBSA by reducing barriers to competition and by providing sufficient training to DoD service acquisition personnel.

In 2000, the DoD formalized its commitment to PBSA. Then-Under Secretary of Defense for Acquisition, Technology, and Logistics Jacques Gansler issued new guidance: "In order to ensure that the DoD continually realizes these savings and performance gains, I establish, at a minimum, that 50% of service acquisitions, measured both in dollars and actions, are to be performance-based by the year 2005."

The *Guidebook for Performance-Based Services Acquisition (PBSA) in the Department of Defense* identifies four elements that are required, at a minimum, for an acquisition to be performance-based: (1) a performance work statement, describing the requirement as a measurable outcome; (2) measurable performance standards, used to define acceptable outcomes and determine if performance thresholds have been achieved; (3) remedies, the incentives and penalties used to provide incentives for performance; and (4) a performance assessment plan detailing performance metrics, as well as how the contractor will be evaluated (Gansler, 2000).

In 2006, the Department of Commerce, Department of Defense, Department of Agriculture, Department of Treasury, General Services Administration, and a private firm, Acquisition Solutions, issued a joint guidebook entitled *Seven Steps to Performance-Based Services Acquisition.*

Trends in PBSA

Within DoD service contracts, the proportion of PBCs versus non-PBCs has changed substantially over the last 15 years. Analysis of the Federal Procurement Data System (FPDS) shows that PBSA now constitutes a majority of total DoD service contract spending. Figure 4 shows the composition of PBCs and non-PBCs among all DoD service contracts. In absolute terms, non-PBCs have declined by more than two-thirds since their peak, from $146 billion in 2008 to $46 billion in 2015, while PBCs have plateaued at approximately $100 billion in 2014 and 2015.

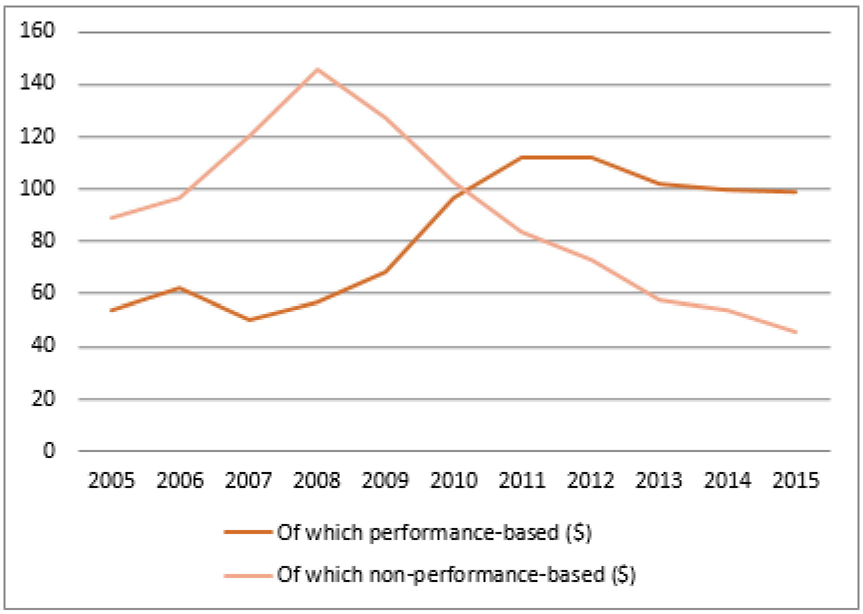


Figure 4. DoD Service Contracts
(Action Obligations in Billions of Dollars; Analysis of FPDS Data)

To a lesser extent, civilian service contracting has seen the same change in contract composition. Performance-based contracts represent more than 50% of all civilian service contracts, and, in 2015, accounted for $70 billion of the total $125 billion spent. Figure 5 shows the composition of service contracts outside of the DoD. The pattern is smoother overall, but shows PBCs overtaking non-PBSA contracts at the same point in time.

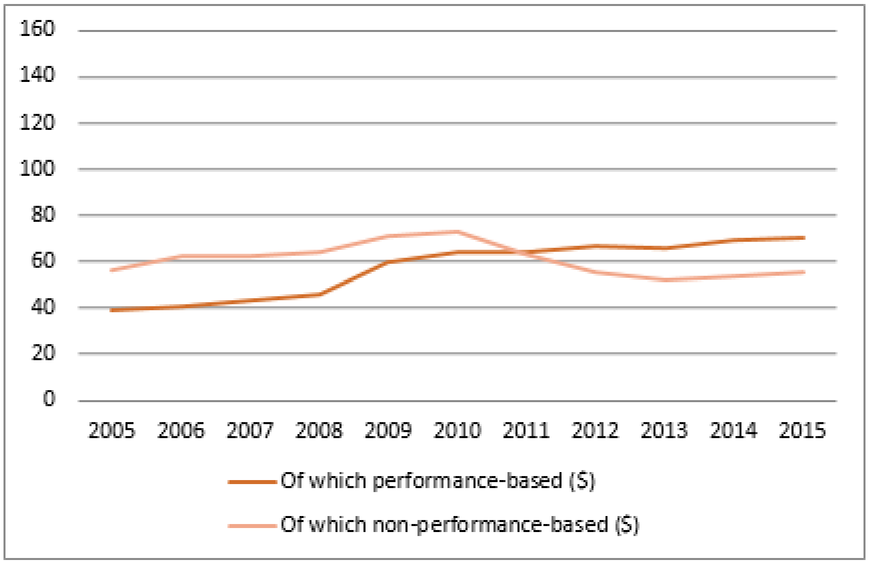


Figure 5. Civilian Service Contracts
(Action Obligations in Billions of Dollars; Analysis of FPDS Data)

It is also noteworthy that as PBSA began to overtake non-PBSA in terms of total service contract spending, so, too, did the DoD begin to overtake civilian government in its rate of PBSA (see Figure 6).

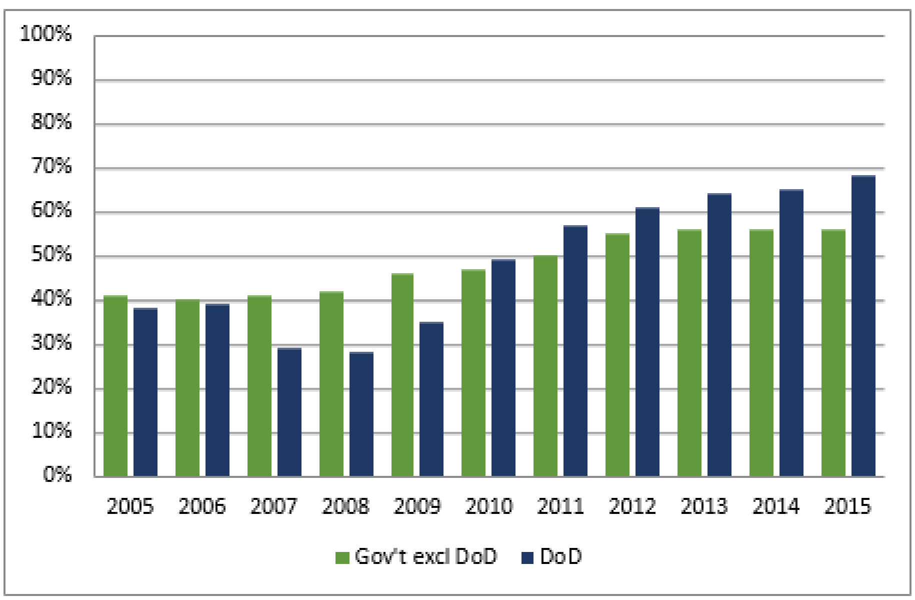


Figure 6. PBSA Contracts
(Action Obligations [$]; Analysis of FPDS Data)

Figure 7 depicts total spending on service contracts over the last 15 years and the percentage of the spending that was performance-based over the last 15 years. The figure also highlights important events that that undoubtedly impacted the use of PBSA. Note that following the 2000 issuance of the directive to increase PBSA such that it would represent 50% of all service contracts spending by 2005, PBSA increased by more than 15% in 2001. The beginning of the War in Iraq saw a continued increase in both services spending and reliance on PBSA. However, between 2006 and 2007 PBSA declined, reaching a four-year low in 2008 (regarded as the height of the war), even as overall contracts spending spiked at over $200 billion.

This decline occurred despite a change in FPDS classification of PBSA contracts. Prior to 2005, FPDS required that “a minimum of 80% of the requirements under the procurement action must meet the FAR standards.” In 2005, the minimum was reduced to 50%. All else equal, one would expect this change to increase PBSA contracts spending. That this was not the case might suggest that the spending figures alone understate the impact of the War in Iraq on PBSA. This decline is unsurprising. Edwards and Nash (2006), who have been critical of PBSA—specifically, its applicability to the provision of complex services—assert that “it is unrealistic to ask agencies to specify services at the time of contract award in clear, specific, objective, and measurable terms when future needs are not fully known or understood, requirements and priorities are expected to change during performance, and the circumstances and conditions of performance are not reliably foreseeable.” There is no doubt that this scenario often prevails during war, which likely explains the apparent reluctance to use PBSA. Interestingly, however, prior research (e.g., Gansler, Lucyshyn, & Rigilano, 2015) indicates that PBSA can be implemented successfully during times of conflict for some types of support, provided that contracts are structured appropriately.

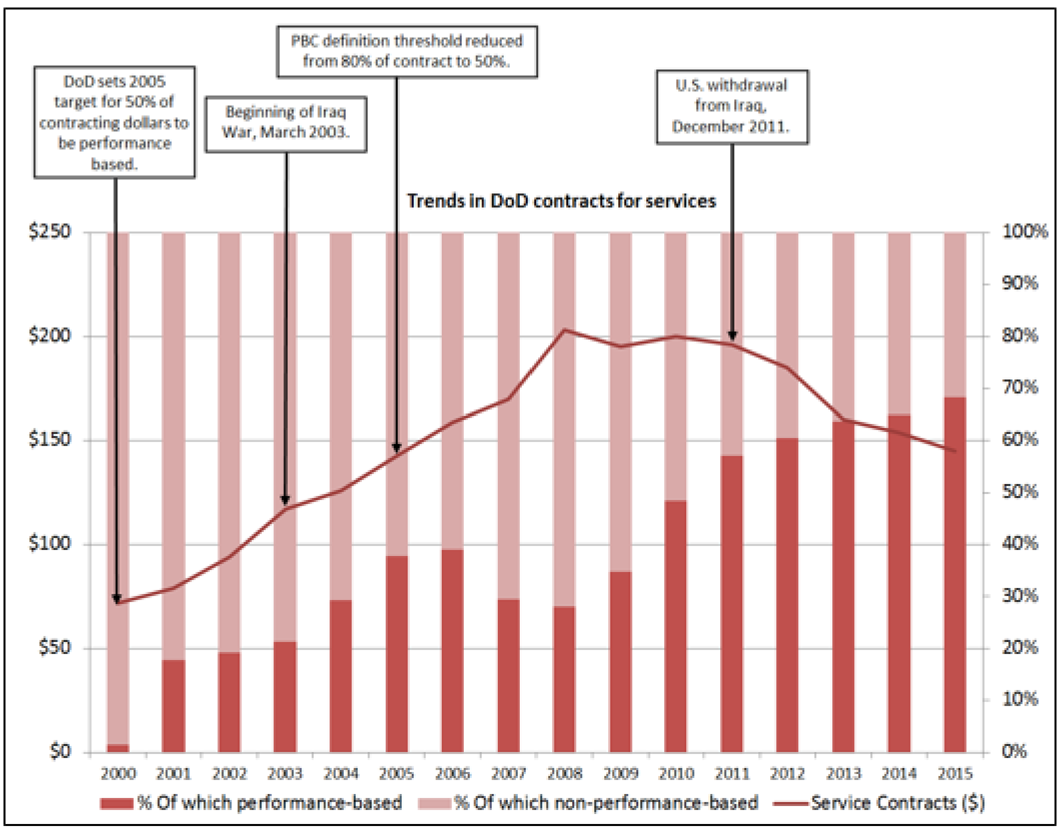


Figure 7. Trends in DoD PBSA
(Action Obligations in Billions of Dollars; Analysis of FPDS Data)

PBSA at Present

PBSA spending has increased dramatically within the DoD since 2000. This increase has been more or less uniform throughout the department. As of FY2015, PBSA rates for the military service branches were as follows: Air Force, 69%; Army, 62%; and Navy, 58%.

Moreover, as previously indicated in Figure 6, the DoD has outpaced civilian government in the use of PBSA (68% versus 56% in FY2015). Of course, this high-level data may not tell the whole story. As indicated previously, the change in threshold from 80% to 50% is not reflected in this data. In addition, some contracts may be performance-based "in name only," either lacking enforcement mechanisms or disbursing payments even when performance is suboptimal. In other words, the data may not accurately reflect the extent to which performance-based strategies are actually applied.

While PBSA may appeal to program officials from a theoretical standpoint, some may be reluctant to embrace this strategy for a variety of reasons, including cultural inertia within the DoD, contractor reluctance, and/or an inability (lack of personnel or technical capacity) to measure contract performance. Indeed, a recent DoD Inspector General report evaluated 60 DoD PBCs. The report revealed that DoD contracting personnel failed to properly negotiate and evaluate most of the contracts. For example, in 33 instances the DoD failed to clearly define criteria for successful completion of various tasks, but disbursed payments to the contractors on a regular basis (DoD IG, 2013).

More generally, a wealth of studies, dating back to the 1980s (e.g., Hart & Holmstrom, 1987) suggest a disinclination on the part of managers to use pay-for-performance strategies for reasons that are "distinctly uneconomic," including notions of fairness, equity, morale, trust, social responsibility, and culture (Baker, Jensen, & Murphy, 1998).

Performance-Based Logistics

Data on one form of PBSA Performance-Based Logistics (PBL) is mixed. The *Defense Acquisition Guidebook* defines PBL as "…the purchase of support as an integrated, affordable, performance package designed to optimize system readiness and meet performance goals for a weapon system through long-term support arrangements with clear lines of authority and responsibility." Application of PBL may be at the system, subsystem, or major assembly level depending on program-unique circumstances and appropriate business case analysis.

While overall PBL expenditure has increased steadily over the last 15 years, likely due to its expansion within successful programs, there were only 87 PBL programs in 2013, compared to over 200 in 2005.
PBL programs evolve along a common trajectory. With new systems, cost-plus reimbursement contracts followed by cost-plus incentive contracts are used in order to provide the government customer and the provider with a cost baseline. Once the costs, risk factors, and system failure modes and rates have stabilized, the program transitions to the use of fixed-price contracts, where providers are paid a fixed cost or fixed rate (e.g., per hour, per mile) so long as operational readiness is achieved at the specified level(s). Over time, the provider makes improvements to its supply chain, logistics networks, operations, and the system itself in order to reduce costs and increase profitability. In the "terminal stage" of its evolution, the exemplary PBL achieves consistently high availability and has optimized maintenance processes and the associated logistics networks on which they rely. The program operates at lower risk, from both a cost and technical perspective.

Despite successful outcomes, there are indications that some longstanding PBL programs are reverting to traditional contracting approaches. Recently, for instance, a high-profile, award-winning PBL program, the High Mobility Artillery Rocket System, transitioned inventory management from the contractor to the government and reverted to cost-reimbursement contracts—as opposed to fixed-price— in an effort to reduce costs. This program is still categorized as "performance-based" in that it relies, at least ostensibly, on performance metrics.

PBSA and Contract Type

However, data on the use of fixed-price PBSA lends support to the supposition that the DoD is increasing its reliance on performance-based strategies (see Figure 8). While it is important to stress that PBSA is a strategy amenable to the use of different types of contracts, fixed-price contracts are generally preferred (FAR 37.102), especially once an acquisition program is well established. Fixed-price contracts incentivize providers to innovate in order to reduce their costs thereby increasing their profit. The cost reductions achieved by the provider can then be taken into account by government in determining baselines for future contracts. Figure 8 indicates that within the DoD, the trend in fixed-price contracting tracks closely with the PBSA trend, both of which overtook "non-performance-based" and "other than fixed-price" in 2010.

At the same time, the figure makes it clear that the fixed-price contracting trend may deviate from the performance-based one. During the height of the Iraq War, in 2008, fixed-price contracts and non-performance-based contracts spiked.

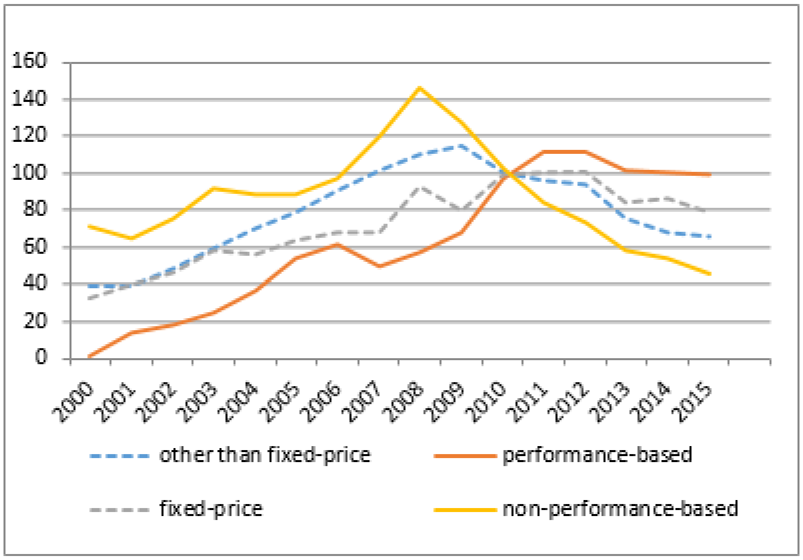


Figure 8. DoD Service Contracts
(Action Obligations in Billions of Dollars; Analysis of FPDS Data)

In fact, between 2006 and 2013, most performance-based contracts were other than fixed-price (see Figure 9). As of FY2015, roughly half of all DoD PBSA was fixed-price. This finding stands in contrast to the relative composition of civilian service contracts, 38% of which were fixed-price as of 2015.

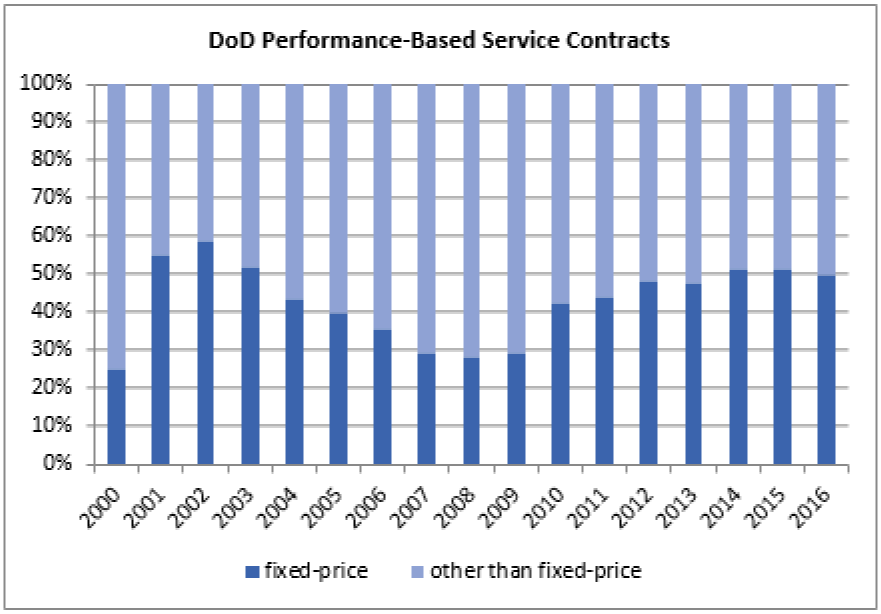


Figure 9. DoD Performance-Based Service Contracts
(Action Obligations [$]; Analysis of FPDS Data)

Case—Navy/Marine Corps Intranet

The Navy-Marine Corps Intranet (NMCI) is a major program that provides information technology services to the United States Navy and Marine Corps. NMCI seeks to provide a streamlined, secure, enterprise-wide network to support the naval shore establishment and connect it with forces at sea by interfacing with the at-sea network. In 2000, the Navy signed a five-year performance-based contract with Electronic Data Systems Corp (EDS) worth an initial $4.1 billion with a three year option to extend. The Navy expected the network to have 412,000 to 416,000 operational computers, or "seats" by FY2004. However, deployment of the network was slower than planned and the program suffered setbacks that delayed its implementation, reduced its desired responsiveness, and increased its cost.

Early in the Navy/Marine Corps Intranet (NMCI) development, the DoN made two important decisions. First, the services of the NMCI would be largely outsourced. Second, the contract would be performance-based.

The DoN primarily sought to contract with the private sector because it did not believe that the DoN had the capability to develop and implement such a holistic information system. Given that the DoN did not believe that it could develop such a capability, and that it wished to implement the NMCI as quickly as possible, contracting much of the technical work to the private sector was the Navy's only realistic option.

The DoN produced an extensive performance plan for the program. The DoN started by identifying its two strategic goals, information superiority and fostering innovation. It then identified nine strategic performance measurement categories and related them to the strategic goals of the NMCI program. These nine categories were interoperability, security and information assurance, workforce capabilities, process improvement, operational performance, service efficiency, customer satisfaction, program management, and network operations and maintenance (GAO, 2006b). The plan included "metrics, targets, and comparative baselines that were to be used for the first annual performance report … [along with the Navy's commitment to] fully develop performance measures for each of the categories and … produce an annual report on NMCI's performance in each of the categories" (GAO, 2006b).

In October of 2000, the DoN awarded the NMCI contract to Electronics Data Systems (EDS). The contract was a "firm-fixed-price, indefinite-delivery/indefinite-quantity contract with performance incentives" (GAO, 2002). The quality of performance was measured according to a set "contractually specified performance level expectations" called service-level agreements (SLAs). The terms of the NMCI contract include delivery and maintenance of workstations and desktop applications; transmission of voice, video, and data; and infrastructure improvements.

The sole-source contract had a five-year base agreement with a minimum value of $4.1 billion, along with a three-year option for an additional $2.8 billion. The contract required delivery of approximately 415,000 seats. The contract was subsequently restructured in 2003 into a seven-year, "$6 billion contract with a three-year option for an additional $2.8 billion" (GAO, 2006b).

The NMCI experienced development difficulties and program revisions early in development initiation. These difficulties became evident once the Navy and the contractor tallied the total number of legacy programs currently operating on Navy and Marine legacy systems. Legacy programs presented numerous compatibility issues. Delays stemmed from the need to (1) undertake an extensive review to list and categorize all legacy applications, (2) develop a new strategy to digest the number of applications that were orders of magnitude larger than originally believed, and, finally, (3) put the new implementation strategy into effect. According to Jordan (2007), "It was initially assumed that the number of these [outdated legacy] applications was in the thousands. After contract award, the Navy and EDS were shocked to find the number was actually 100,000." The contract goal of transitioning legacy applications into 500 NMCI accredited programs was revealed to be a much more difficult task than initially thought.

By May 2002, only 4,000 seats had been cutover. Due to NMCI's slow progress, Congress, in December 2002, sought to strengthen oversight by requiring authorization to increase the seat limits beyond 60,000, and then up to 150,000.

In 2003, EDS shareholders filed a class-action lawsuit against the company alleging security fraud stemming from second quarter losses, primarily due to "problem contracts." According to EDS, difficulties with the NMCI contract resulted in a $334 million pretax loss on the program through 2003 (Verton, 2003). Subsequently, the DoN and EDS restructured the NMCI's contract and implementation schedule. One report estimates that EDS losses averaged $800 million annually in the first years of the contract, totaling $3 billion (Jordan, 2007).

Acknowledging the NMCI's shortcomings , the Navy awarded a one-year $5.9 million contract to BearingPoint in December 2006 to help manage IT services (Beizer, 2006). BearingPoint was awarded a larger five-year contract, with a maximum value of $57.9 million in October 2007, principally to "design and operate a secure, battle-ready global information technology network for the Naval Network Warfare Command" (Hubler, 2007). This action solidified the subtle—if unofficial—shift away from the NMCI's initial goal of information superiority (in the form of a battle-ready information system) to simply furnishing the DoN with an operational information network.

The NMCI experience demonstrates that firm-fixed-price contracts for high-risk, ambitious programs do not necessarily reduce program costs. Rather, fixed-price contracts are ideal when requirements are known and stable, and the technical risk is low. The experience also shows that the metrics included in PBCs may produce unfavorable outcomes if consequences are not anticipated. For instance, the metrics involving e-mail transfers and the percent of bandwidth used to provide connection to external networks provided EDS an incentive to severely limit the size of e-mail attachments, frustrating many who were unable to transmit larger files.

Following the program's early challenges, NMCI steadily improved. On September 30, 2010, the NMCI contract ended and the new Continuity of Services Contract (COSC) began. Today, NMCI is one of the largest intranets in the world, providing end-to-end secure IT services to more than 400,000 computers and 800,000 users across 2,500 locations (DoN, 2017). At present, the Navy is transitioning NMCI services to the Next Generation Enterprise Network (NGEN). The NGEN acquisition approach will allow NMCI to transition from a "monolithic model" to a segmented business model that allows for periodic competition of segmented services (DoN, 2017).

Case—Stryker PBL

The Army's first new vehicle since the early 1990s, the M1126 Stryker, is a rapidly-deployable, wheeled armored vehicle. Stryker successfully combines resiliency, mobility, and versatility, creating the ideal combat vehicle, and quickly becoming an essential tool for the United States Armed Forces.

As top Army officials became increasingly frustrated with the attributes of existing combat vehicles—many of which were either too heavy to be deployed efficiently or too light to be effective in combat—the Army began its search for a new armored vehicle for its fleet. The acquisition process was accelerated as U.S. troops in Iraq and Afghanistan began to encounter unprecedented threats from improvised explosive devices (IEDs), heightening the need for a new armored vehicle. Indeed, the Stryker vehicle was among the fastest acquisitions of a major weapons system in U.S. history.

Awarded to General Dynamics Land Systems (GDLS) in 2000, the initial PBL for the Stryker vehicle fell under a cost-plus-fixed-fee (CPFF) portion of a larger contract for vehicle manufacture and delivery. The CPFF covered "all fielded vehicles in garrison or deployed" (Coryell, 2007) for a five year period. Under the agreement, GDLS would produce and repair and maintain vehicles at four primary locations: Anniston, Alaska, Ontario, and London (Denizer, 2007).

The CPFF contract was chosen to provide maximum flexibility to meet rapidly-evolving conditions while allowing Army officials to gauge the costs associated with different levels of performance so that a firm-fixed-price (FFP) contract could be used at a later point (Coryell, 2007). The initial contract specified a single metric, an operational readiness rate (ORR). Vehicles were expected to meet a 98% ORR during fielding and training exercises, at least a 90% ORR for deployed vehicles.

GDLS would go on to win two follow-on PBL contracts, in 2007 and 2012, valued at $1.5 billion and $2.5 billion respectively, which would extend sustainment and support for the growing Stryker fleet. In 2007, some 1,500 vehicles had been fielded; by 2012, this number had risen to close to 2,500 (DoD IG, 2012).

For the first two Stryker brigades that deployed to Iraq, Army officials reported ORRs averaging 96% from October 2003 through September 2005 (GAO, 2006a). In addition, the Army consistently noted that contractors were providing impressive levels of support, and according to a 2006 GAO report, they were more knowledgeable and efficient than their military counterparts with regard to the specifics of the Stryker vehicles (GAO, 2006a). The program's use of contractor personnel for sustainment efforts allowed soldiers on the ground to participate in extra trainings and perform other necessary, military specific roles (GAO, 2006a). Pre-existing relationships between soldiers within SBCT and deployed contractors also created a successful and effective work environment overseas.

From a cost perspective, however, contract performance is less clear. In 2012, the DoD Inspector General asserted that the follow-on contract's continued use of a sole metric (readiness) in combination with a high-ceiling, cost-plus contract unduly incentivized the contractor to accumulate significant excess inventory valued at $335.9 million (DoD IG, 2012). The Army responded that the excess inventory could be attributed, in part, to contractor improvements in reliability, and that the spare parts would be used eventually, albeit at a slower pace than anticipated (DoD IG, 2012).

Given the Army's heavy reliance on Stryker during the Iraq War, changing operational tempos, and the lack of historical cost data, the use of a cost-plus-fixed-fee contract (as opposed to a fixed-price contract) was well-founded. However, it appears that the Army could have implemented better cost controls, perhaps by tying the fixed fee to an agreed-upon cost-per-mile metric. As indicated in the previous section, the DAU lists cost-per-unit metrics as essential indicators of PBL performance.

Findings

The following is a summary of our findings:

- Based on the available data, the DoD has made impressive gains in its implementation of PBSA. In 2016, close to 70% of DoD service contracts were performance-based.
- The rate of PBSA within the DoD has increased steadily since 2010, even as overall spending on services has decreased.
- Despite increases in the overall rate of PBSA, PBL implementation, in terms of the number of programs, has declined.
- The DoD has outpaced the rest of the government in the implementation of PBSA.
- DoD guidance states that fixed-price contracts are preferred within the context of PBSA. The proportion of DoD performance-based contracts that are fixed-price has increased to approximately 50% in 2016, up from a low of 29% in 2007. The civilian PBSA rate was 38% in 2016.
- PBSA is not the right choice for all acquisitions and even when it is the right choice, performance-based contracts are not always structured appropriately.
- Implementing PBSA, and developing the appropriate metrics is more challenging than traditional contracting, and requires a different skillset.

Recommendations

Based on these findings, we provide the following recommendations:

- **DoD Senior Leadership should continue to emphasize the use of PBSA.** Successful reform—especially transformation of a bureaucratic culture—takes concerted effort over a prolonged period of time. Top-level management must lead this reform to produce "buy-in" at lower levels. To be effective, leadership must continuously communicate its vision and support its message. Leadership should reaffirm its commitment to PBSA by issuing memoranda that stress the importance of using PBSA. Leadership should follow up initial support by periodically issuing memoranda that update the DoD on the use of PBSA and its success stories.

- **The DoD should ensure proper alignment of government objectives with provider incentives**.
  PBSA arrangements can be more challenging to develop and manage than other contract types. Just as an appropriate program structure aligns the incentives of the customer (the government) and the support provider, leading to a win-win scenario, an inappropriate structure can create perverse incentives, and result in undesired or unintended consequences.
- **The DoD should improve the training of the acquisition workforce.**
  The DoD should also increase the training of its employees involved in the acquisition of services. Training should emphasize the importance of a robust requirements definition process, the need for clear performance requirements, measurable performance, and standards.
- **Stringent cost controls must be applied to cost-plus performance-based service contracts.**
  Categorizing a contract as performance-based does not make it so, especially with regard to cost-plus contracts. While some performance-based service acquisitions are best suited to cost-plus contracts, they must be structured appropriately to ensure optimal outcomes. Carefully considered contract ceilings, cost-per-unit usage rates, and logistics footprint constraints should be included in cost-plus contracts. Without these features, contractors may be incentivized to accrue surplus inventory beyond what is necessary to meet the performance requirement.

References

Beizer, D. (2006). BearingPoint to help steer NMCI into safe berth. Retrieved from http://washingtontechnology.com/articles/2006/12/13/bearingpoint-to-help-steer-nmci-into-safe-berth.aspx

Center for Medicare and Medicaid Services. (2015). *Hospital value-based purchasing.* Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/index.html?redirect=/hospital-value-based-purchasing/

Denizer, G. (2007, June). *Applicability analysis of performance based logistics implementation for the U.S. Army Stryker Armored Vehicle to improve Turkish Army weapon system support.* Retrieved from http://calhoun.nps.edu/handle/10945/10244

Dennis, M., & Kambil, A. (2003). Service management: Building profits after the sale. *Supply Chain Management Review, 7*(1), 42–48.

DoD Inspector General (DoD IG). (2012, November 30). *Accountability was missing for government property procured on the Army's services contract for logistics support of Stryker vehicles.* Retrieved from http://www.dodig.mil/pubs/documents/DODIG-2013-025.pdf

DoN. (2017). Next generation enterprise. Retrieved from http://www.public.navy.mil/spawar/PEOEIS/NEN/NGEN/Pages/2013Announcement.aspx

Edwards, V. J., & Nash, R. C., Jr. (2006, August). Relational contracting: A proposal for a new approach to performance-based services acquisition. *Contract Management,* 32–

40. Retrieved from http://www.ncmahq.org/docs/default-source/default-document-library/classificationspeciman/01_ae7b7_cm0806_f5

Federal Acquisition Regulation (FAR), 48 C.F.R. 2.1-13 (2016).

Gansler, J. (2000). *Guidebook for performance-based services acquisition (PBSA) in the Department of Defense*. Retrieved from http://www.acq.osd.mil/dpap/Docs/pbsaguide010201.pdf

GAO. (2006a). *Defense logistics: Changes to Stryker vehicle maintenance support should identify strategies for addressing implementation challenges*. Washington, DC: Author.

GAO. (2006b). *Information technology: DOD needs to ensure that Navy Marine Corps Intranet program is meeting goals and satisfying customers*. Washington, DC: Author.

GAO. (2008). *Defense logistics: Improved analysis and cost data needed to evaluate the cost-effectiveness of performance based logistics.* Washington, DC: Author.

GAO. (2009). *Defense acquisitions: Actions needed to ensure value for service contracts*. Washington, DC: Author.

Hubler, D. (2007). BearingPoint to upgrade Navy IT network. Retrieved from http://washingtontechnology.com/articles/2007/10/03/bearingpoint-to-upgrade-navy-it-network.aspx

Jordan, K. (2007). *The NMCI experience and lessons learned: The consolidation of networks by outsourcing*. Retrieved from http://ctnsp.dodlive.mil/files/2006/12/Case-12-The-NMCI-Experience-and-Lessons-Learned.pdf

Martin, L. (2016). Performance-based contracting. In R. A. Shick (Ed.), *Government contracting: A public solutions handbook* (pp. 61–74). New York, NY: Routledge.

National Institute of Governmental Purchasings (NIGP). (2012). *Performance based contracting*. Retrieved from https://www.nigp.org/docs/default-source/New-Site/global-best-practices/performancebased.pdf?sfvrsn=2

Taylor, K., & Shaver, M. (2010). Performance based contracting: Aligning incentives with outcomes to produce results. In M. F. Testa & J. Poertner (Eds.), *Fostering accountability* (pp. 291–327). Oxford University Press.

Verton, D. (2003). EDS files motion to dismiss shareholder lawsuit. Retrieved from http://www.computerworld.com/s/article/86397/EDS_files_motion_to_dismiss_shareholder_law

Acknowledgments

This research was partially sponsored by a grant from the Naval Postgraduate School.

FAR-Based Crowd Sourcing

Commander Zachary Staples, USN—is the Director of the Center for Cyber Warfare at the Naval Postgraduate School (NPS) and a Military Associate Professor of Electrical and Computer Engineering at NPS, located in Monterey, CA. [zhstaple@nps.edu]

Abstract

Looking to the past, our Navy acquisition process was organized to deliver advanced industrial hardware that has resulted in the most powerful fleet in the world. Moving forward we will still need world-leading hardware, but now and into the future warfighting capability is generated by software, perhaps more than hardware. Yet acquisition processes and timelines for the delivery of software and hardware may not necessarily be the same. This talk compares and contrasts exemplars of world-leading hardware and software companies against the pace of innovation and acquisition in the Navy, and then goes on to demonstrate a groundbreaking example of global crowdsourcing to achieve both cost and time savings on a maritime domain specific software development project.

Talk Summary

The Chief of Naval Operations, Admiral John Richardson, uses a synonym for speed in almost every speech he gives. Whether he is emphasizing the pace of strategic change or the need for high velocity learning, there is an urgency to his message that we must think in a competitive framework and acquire the capability to fight in a more rapid and accelerated manner (Freedberg, 2017; Richardson, 2017).

As we move to a more agile acquisition mindset, it is helpful to find benchmarks for complex systems being manufactured around the globe and compare the U.S. Navy's speed of development and delivery. Interestingly, some of the most advanced commercial organizations in the world involved in producing high-tech hardware are constrained to an approximate annual model year cycle. Of note, the iPhone was introduced by Apple in 2007 and 10 years later, in 2017, they have produced only 11 distinct models (iPhone, 2017). Similarly, Tesla is often regarded as one of the most innovative and creative technology manufacturers in the world, but since their inception they have not produced vehicles faster than a model year pace (Tesla, 2017).

Thus, it is illuminating to consider that the U.S. Navy and its contractors have produced approximately two large surface combatants and two fast attack submarines per year for the past several years and are expected to continue this pace into the future (Office of the Chief of Naval Operations, 2016). Due to technology insertion programs in both the surface and submarine forces, these ships represent model year changes from year-to-year, keeping pace with the very best hardware producers in the world.

In stark contrast to the pace of high-technology hardware, leading software corporations have created development operations (dev-ops) processes that give them the capability to deliver production code multiple times per day. For example,

> At Facebook, code can be released twice a day, but this is done mostly for bug fixes and internal code. New production code is released once per week: thousands of changes by hundreds of developers are packaged up by their small release team on Sundays, run through automated regression testing, and released on Tuesday if the developers who contributed the changes are present. (Bird, 2013)

Bear in mind that Facebook is delivering mission and revenue critical software upgrades and patch fixes to over a billion users around the world, and they are deploying code onto a hardware constellation that includes tens of thousands of device configurations that they have no control over. Yet, every week our handheld devices, tablets, laptops, and desktop computers run the new Facebook code with rare complications.

Similarly, the software branch at Tesla has a patch push infrastructure, and in September 2016 patched a flaw that was exposed on YouTube in one day and had updated the entire Tesla fleet of model-S vehicles in 10 days (Reuters, 2016).

In stark contrast to the world-leading commercial dev-ops capability demonstrated by Facebook and Tesla, consider the multi-year development of iterations to the AEGIS weapons system, or the much publicized “millions of lines of code” that delayed the Joint Strike Fighter for years. While the Navy is keeping pace with world-leading hardware manufacturers, our CNO calls for accelerated acquisition because we are falling years behind the pace of world-leading software firms.

If the Navy needs innovation and speed in our software acquisition, I chose to look at a few historical examples to inspire a path to change in the present. First, consider the 1927 Orteig Prize, design to be awarded to the first allied aviator to fly non-stop from New York to Paris or vice versa. When Charles Lindbergh was catapulted to fame by his successful claim on the prize, it was the culmination of a widespread investment in aviation technology that was set in motion by the drama and promise of the prize (Williams, 2015).

Similarly in 1714, the British government offered a reward to the man who could find the longitude of a location at sea. Before the discovery of a means to measure longitude, ships would drift off course or get lost at sea because they were unable to pinpoint their location relative to their east-west travel. Drawn to the incentive of a prize, clockmaker John Harrison created what is known today as the chronometer, and astronomer Tobias Mayer perfected the astronomic tables that refined the efficacy of Harrison’s machine and longitude was solved (Dunn & Twigg, n.d.).

Following in the footsteps of these grand prizes, this research sought to use a prize, awarded via a Federal Acquisition Regulation (FAR) compliant contract, to inspire speed and innovation in a Navy-specific software application.

Thankfully, in the wake of global communities brought together by the Internet there are several commercial organizations specializing in online competitions motivated by prize money. Some examples include Topcoder.com, HeroX.com, and Kaggle.com. Each of these platforms has a community with specific skills, and many individuals that form the community compete to earn prizes as their full-time employment.

An example of a software problem that my research team deals with every day has to do with creating radio frequency scenarios to stimulate a game system I deploy in the Navy hackathons I run. So we challenged the online community at Topcoder to design a user interface to solve the time lag with manually plotting vessels (Topcoder, 2016).

The crowdsourced challenge received a huge response, with 61 submissions from 17 different countries. To my knowledge this is the Navy’s first use of a globally crowdsourced online competition to develop software for a maritime task.

The Navy usually solves problems like this by hiring a vendor for a specific long-duration code development process. By using a contest to attract the help of the best software engineers in the world, I finished the hackathon with three world-class, Navy-owned designs for the price of $12,000 dollars and created in just three days.

The contract to deliver this software was awarded to a single vendor for services and products. The services of the vendor included managing the crowdsourced development on Topcoder, translating the military requirements into a community challenge, and providing real-time management of the community's questions during the event. The deliverable products were the three designs.

After repeatedly hearing about acquisition reform but seeing limited results, it is time to accept the current acquisition system and start finding new ways to advance military software on a timeline that differs from hardware acquisition. This process should take full advantage of existing controls in the FAR, but use prize-based contests to attract the best coders and programmers in the world. It is possible to deliver world-class software at a fraction of the cost and time required from traditional defense partners, and this pathway is ready for more serious exploration on projects of greater significance than our lab tool.

References

Bird, J. (2013). This is how Facebook develops and deploys software. Should you care? Retrieved from https://dzone.com/articles/how-facebook-develops-and

Dunn, R., & Twigg, S. (n.d.). Longitude Prize: The history. Retrieved from https://longitudeprize.org/history

Freedberg, S. J. (2017). CNO Richardson urges fast-track for cyber, EW & drones. Retrieved from http://breakingdefense.com/2015/12/cno-richardson-urges-fast-track-for-cyber-ew-drones/

iPhone. (2017, April 5). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/iPhone

Office of the Chief of Naval Operations. (2016). *Report to Congress on the long-range plan for construction of naval vessels for fiscal year 2017*. Washington, DC: Author.

Reuters. (2016). Tesla fixes security bug after claims of Model S "white hat" hack. Retrieved from http://fortune.com/2016/09/20/tesla-security-bug-hack/

Richardson, J. (2017). Soun*dings—CNO talks directed energy, competition, and faster acquisition* [Video file]. Retrieved from https://soundcloud.com/cnorichardson/soundings-cno-talks-directed-energy-competition-and-faster-acquisition

Tesla. (2017). Tesla's unfathomable launch. *Bloomberg*. Retrieved from https://www.bloomberg.com/news/articles/2017-03-27/tesla-model-3-ramp-up-aims-to-crush-bmw-and-mercedes

Topcoder. (2016). RUX—*Boat in a box simulation app design concepts challenge*. Retrieved from https://www.topcoder.com/challenge-details/30056506/?type=design

Williams, M. (2015). History of challenges: The Orteig Prize (1919–1927). Retrieved from https://herox.com/news/428-history-of-challenges-the-orteig-prize-1919-1927

THIS PAGE INTENTIONALLY LEFT BLANK

Panel 20. Trends in International Defense Acquisition

Thursday, April 27, 2017	
3:30 p.m. – 5:00 p.m.	**Chair: Francois Melese**, Professor, Defense Resources Management Institute **Discussant: Laura Armey,** Assistant Professor, Defense Resources Management Institute ***Strengths and Weaknesses of China's Defense Industry and Acquisition System and Implications for the United States*** Tai Ming Cheung, University of California San Diego ***A Systemic Analysis of Military Equipment Acquisition Among NATO Suppliers: A Proof of Concept Based on a Multi-Layered DSS Approach*** Martin Zsifkovits, University of the Federal Armed Forces Munich, Germany Gonzalo Barbeito, University of the Federal Armed Forces Munich, Germany Dieter Budde, Major General (Ret.), German Armed Forces Max Krüger, University of Applied Sciences Furtwangen, Germany Stefan Pickl, University of the Federal Armed Forces Munich, Germany

Francois Melese—Professor, Naval Postgraduate School, received his BA in Economics from the University of California at Berkeley in 1977, MA in Economics from the University of British Columbia, Canada, in 1979, and PhD from the University of Louvain, Belgium, in 1982. He was previously a Research Fellow at the Institut de Recherches Economiques et Sociale (IRES), University of Louvain, Belgium, and Assistant Professor of Economics at Auburn University. Dr. Melese has papers published in the *Quarterly Journal of Economics*, the *Southern Economic Journal*, *Energy Economics*, the *International Trade Journal*, and *Defense Analysis*. He has presented papers at meetings of the American Economic Association, the European Economic Association, the Southern Economic Association, and the World Econometric Society, as well as at meetings of the International Association of Science and Technology for Development. He is a member of the American Economic Association, Southern Economic Association, Operations Research Society, and the Research Society of American Scientists–Sigma XI. Dr. Melese joined the faculty of DRMI in June 1987.

Laura Armey—Assistant Professor, Naval Postgraduate School, received a BA in Anthropology (2004), a BA and MA in Economics (2004), and a PhD in Political Economy and Public Policy (2008) from the University of Southern California. Dr. Armey previously worked as an analyst at the Cost Analysis and Program Evaluation division of the Office of the Secretary Defense, where she provided economic analysis for defense policy making. One area of her research focuses on the political and economic factors that affect post-war reconstruction and other outcomes of conflicts. An additional area of her research focuses on the impact of combat exposure on the economic outcomes of service members. In addition, she has worked on analysis of telecommunications markets and the impact of communication technology in developing countries, both at DRMI and as a post-doctoral research fellow for the Institute for Communications Technology at the University of Southern California. Dr. Armey joined the DRMI faculty in August 2010.

Strengths and Weaknesses of China's Defense Industry and Acquisition System and Implications for the United States[1]

Tai Ming Cheung—is Director of the Institute on Global Conflict and Cooperation located at the University of California, San Diego, in La Jolla. He leads the institute's Study of Innovation and Technology in China project that examines China's efforts to become a world-class science and technology power. Dr. Cheung is also an associate professor at the School of Global Policy and Strategy at UC San Diego, where he teaches courses on Chinese foreign and defense policy and Chinese security and technology policy. Dr. Cheung has a PhD in War Studies from King's College London. [tcheung@ucsd.edu]

Abstract

Major technological progress is taking place across virtually the entire spectrum of the Chinese defense industry, from traditional sectors such as aerospace and seapower to the newer domains of space, information technology, and cyber. This is steadily narrowing the defense technological gap with the United States. At the same time though, the Chinese defense industry and acquisition system is plagued by deep-seated problems that call into question whether the current progress is sustainable over the long term. Understanding the state, reforms, and prospects for China's defense industry and acquisition system is of critical importance to the United States. This paper examines the dueling realities of China's defense industry and acquisition system. A central reason that the Chinese defense industry has been able to be keep costs down and accelerate the pace of acquisition is because it has operated on an absorption-based, good-enough development model. But as it transitions to more of an original innovation-based, higher end development framework, risks grow significantly, and this will impact the costs and pace of the acquisition process

Introduction

The opening decades of the 21st century have been a gilded age for China's defense industry and acquisition system. Lavished with high-level leadership attention and ample funding, the country's armament community from scientists to testers have been busy on a scale not seen since the Soviet–U.S. Cold War in the second half of the 20th century. Major progress is taking place across virtually the entire spectrum of the Chinese defense industry, from traditional sectors such as aerospace and seapower to the newer domains of space, information technology, and cyber. This is steadily narrowing the defense technological gap with the United States.

At the same time though, the Chinese defense industry and acquisition system is plagued by deep-seated problems that call into question whether the current progress is sustainable over the long term. The chief defense acquisition tsar, General Zhang Youxia, director of the Central Military Commission Armament Development Department (CADD), highlighted this predicament in a speech in 2014 when he said that structural and process problems have become such an obstacle that the primary "bottleneck issue for armament development is no longer the shortage of funds or technology. Instead, institutional systems and mechanisms have become the greatest hurdle to armament building and development"

---

[1] Draft version: Do not quote or distribute without author's permission.

(Zhang, 2014). Zhang added that if these impediments cannot be removed, future progress in weapons development “may just be empty talk.” Since then, the Chinese authorities have launched major reform initiatives to tackle these problems.

Understanding the state, reforms, and prospects for China’s defense industry and acquisition system is of critical importance to the United States, countries in the Asia-Pacific, and the international community because of China’s rise as a global power. A key enabler of China’s growing might and external influence is its military technological capabilities, as evidenced by the acquisition of long-range power projection and anti-access/area denial capabilities. For the U.S, China represents a “pacing threat” and is its chief long-term defense technological competitor. China’s technological transformation is one of the principal drivers behind the U.S. Defense Department’s Third Offset Strategy aimed at addressing the long-term erosion of U.S. defense technological superiority.

This paper examines the dueling realities of China’s defense industry and acquisition system. The starting point is an overview of the nature and characteristics of the Chinese defense acquisition system. Attention will then turn to assessing the causes behind the rapid progress that has been made within the past two decades in the development of China’s defense science and technology capabilities and the role that the acquisition system has played. The focus will then turn to analyzing the structural constraints and weaknesses, especially as they relate to the defense acquisition system, that could pose serious barriers to the China’s continuing defense technological and industrial advancement. The paper concludes with a discussion of the implications for the United States.

The Nature of the Chinese Defense Acquisition System

Although the People’s Liberation Army (PLA) in coordination with the Chinese defense industry has been engaged in the design, engineering, test and evaluation, production, and operation and support of defense systems since the 1950s, acquisition was considered of lower-order importance compared to warfighting, political discipline, and logistics until the end of the 1990s. It was not until 1998 that the PLA high command established a separate acquisition organization called the General Armament Department (GAD) to serve alongside the General Staff Department (GSD), General Political Department, and General Logistics Department.

Previously, the acquisition system had been divided between two competing systems. One portion was manned by serving military officers who worked within the GSD as well as in armament management entities at the service-level, while another component comprised civilians and uniformed personnel who were attached to the Commission for Science, Technology, and Industry for National Defense (COSTIND) that was a state entity and represented the interests of the civilian defense industry. This bifurcated arrangement was a product of the Socialist central planning system that managed the Chinese economy between the 1950s until the 1990s. It meant that the acquisition system was compartmentalized between civilian and military masters, which led to constant and often bitter bureaucratic infighting because these two groups had widely divergent interests. As the consumer, the military wanted weapons that could be produced on time, met its specifications, and were cost-effective. But the defense industry had little incentive to meet the PLA’s requirements because it faced little competition.

The establishment of the GAD led to a far-reaching reorganization of the acquisition system. A crucial change was that the GAD assumed primary responsibility for acquisition matters, while the defense industry—through COSTIND—was relegated to a supporting role. This change did not happen overnight though and was met with considerable resistance because of the entrenched control that the defense industry had enjoyed over the acquisition process for several decades. But the GAD gradually consolidated its management of the acquisition system until it was replaced by the CADD at the beginning of 2016.

A defining characteristic of the current Chinese defense acquisition system is that it operates as a classic command and control regulatory system. The military authorities rely on administrative coercion and threats to achieve compliance, they are responsible for direct micro-management and rule-making, and the primary focus of rules and regulations is more on addressing what enterprises do rather than on their performance or outputs (Ogus, 2002, pp, 21–22). While this top-down regulatory approach may work in a centrally planned system, it is poorly suited to market-based environments. One major reason is that activity-based intervention requires regulatory agencies to have adequate information to monitor what is going on, but their access to data is limited in more open markets because they are required to play a less direct regulatory role and allow enterprises greater autonomy and privacy.

The model used in the United States and other advanced economies is the independent regulatory system that emphasizes the importance of political independence, impartiality, and transparency. The Chinese defense industry began to take concrete steps towards the establishment of a more independent regulatory structure in the late 1990s, especially taking advantage of reforms to separate the military-civilian regulatory system by rolling back the authority and intervention of COSTIND and allowing defense corporations to have greater autonomy. COSTIND's status and influence was further reduced in 2008 when it was demoted from a super-ministry to a sub-ministerial agency under the Ministry of Industry and Information Technology (MIIT) and retitled as the State Administration for Science, Technology, and Industry for National Defense (SASTIND; Mao, 2012, pp. 58–60).

Another distinguishing feature of the Chinese defense acquisition system is that it operates a predominately absorptive model of technology development. This is typical of catch-up countries whose domestic research and development capabilities still lag behind the world's advanced defense technology powers. Absorption-oriented acquisition systems organize and operate in a fundamentally different way from innovation-based systems like the United States'. Two differences stand out. First, absorption is a low-risk, high-reward enterprise because the technological development path has already been mapped out. Second, absorptive systems place overwhelming priority in investing in engineering, capabilities, especially related to reverse engineering, and less on research and development.

The primary benefits from absorption are significant cost savings and time reductions. This has allowed the Chinese defense establishment to narrow, and in a few cases eliminate, the technological gap with regional and global competitors in an expanding number of areas. The biggest beneficiaries have been in the aviation, naval shipbuilding, and select precision strike missile sectors. Without these technological achievements that are being translated into operational capabilities, the PLA's shift to a more regionally assertive and maritime-oriented posture would have been little more than empty talk.

The military aviation sector has been especially reliant on the leveraging of foreign technology transfers to support its development.[2] A significant proportion of the country's combat aircraft development programs have depended on foreign, mostly Russian, technology inputs. These technology transfers come in several forms:

- **Reverse engineering:** The Chinese aviation industry has been able to reverse engineer complete platforms acquired through license assembly agreements (Su-27), off-the-shelf purchases (Su-30MK2), and opportunistic acquisition of prototypes (Su-33), which it then adapts and indigenizes with local sub-systems and components. A substantial proportion of the PLA Air Force's combat inventory consists of these re-innovated aircraft, such as the J-11B (Su-27), J-15 (Su-33), and J-16 (Su-30MK2).
- **Research and development assistance:** A number of Chinese "indigenous" programs have received extensive levels of foreign assistance in their design and development, including co-design and co-development. Much of the original design of the J-10A fighter, for example, was from Israel. China and Russia are also reportedly close to signing an agreement on the co-design and development of a heavy-lift helicopter, of which Russia is in charge of aerodynamics design and China providing avionics systems ("China, Russia to Co-Develop," 2015).
- **Critical components and sub-systems:** While the overall technological level of the Chinese aviation industry is steadily improving, there are pockets of backwardness in critical components and sub-systems. High-end turbofan jet engines stand out as the biggest weakness, which has made China dependent on Russian engines.
- **Enabling technologies:** As the Chinese aviation industry becomes more sophisticated across all the stages of the research, development, and acquisition process, it is sourcing foreign assistance for wind tunnels, computer-aided design and manufacturing software, and advanced production equipment such as multi-axis machine tools.

Foreign technology has also played an influential role in the improving technological performance of the naval shipbuilding industry. This was especially the case from the 1990s to the mid-2000s when there was extensive importation of Russian technology and knowhow. As Chinese shipbuilders absorbed these transfers, they have been able to substantially reduce their foreign reliance in the past decade. A 2015 U.S. Office of Naval Intelligence (ONI) assessment of the equipment modernization of the PLA Navy's surface fleet noted that, "By the second decade of the 2000s, the PLA(N)'s surface production shifted to platforms using wholly Chinese designs and that were primarily equipped with Chinese weapons and sensors (though some engineering components and subsystems remain imported or license produced in country)" (ONI, 2015, p. 14).

The ONI report noted that the last purchase of a foreign naval platform was the Sovremennyy-II guided missile destroyer in 2006 and since then, the Chinese naval shipbuilding industry has been engaged in

[2] For useful background analysis, see Saunders and Wiseman (2011).

much longer production runs of its domestically produced surface combatants and conventional submarines, suggesting greater satisfaction with recent designs. The Jiangkai-class (Type 054A) frigate series, Luyang-class (Type 052B/C/D) guided missile destroyer (DDG) series, and the upcoming new cruiser (Type 055) class are considered to be modern and capable designs that are comparable in many respects to the most modern Western warships. (ONI, 2015)

Faster and Cheaper: China's Accelerated Defense Acquisition Process and Comparisons With U.S. Accelerated Programs

Taking advantage of the absorptive low-risk, high-reward development model, the Chinese defense acquisition community has been able to successfully undertake the rapid development, production, and deployment of select high priority weapons projects over the past couple of decades. The speed of China's achievements initially surprised outside observers, especially in the late 2000s and early 2010s when a number of Chinese programs were unveiled.[3] This is because of the checkered historical track record of the Chinese defense industry, especially before the late 1990s, in which many projects failed or were seriously delayed because of bureaucratic red tape, lack of support, or inadequate resources.

The Chinese accelerated acquisition process has some of the following features:

- **Concurrent development, testing, and low rate initial production:** This sees the compression, overlapping, or even skipping of various phases in the acquisition process with the goal of getting programs into production and deployment as quickly as possible. Some of this compression occurs with concurrent technology maturation, risk reduction, and development as well as concurrent production and deployment.
- **Accelerated research & engineering development, but often lengthy delays in early production phases:** A number of Chinese weapon development programs have been rushed through the initial research and development phases, but then spend extended periods of time undergoing prototyping or demonstration testing. If compared to the U.S. acquisition process, this would be the equivalent of manufacturing readiness levels (MRL) 5 (capability to produce prototype components in a production relevant environment) and 8 (pilot line capability demonstrated; ready to begin low rate initial production).
- **High-level leadership attention and active intervention:** The Chinese authorities have designated the development of a small hand-picked number of strategic weapons and technology capabilities to be critical national priority and have established oversight mechanisms to allow senior top-level national leaders such as the Communist Party General Secretary, who is concurrently

[3] For example, Vice Admiral David Dorsett, deputy chief of Naval Operations for Information Dominance, said in 2011 that the U.S. intelligence community has "been pretty consistent in underestimating the delivery and IOC [initial operational capability] of Chinese technology, weapon systems. They've entered operational capability quicker [than expected]."

the commander-in-chief, and prime minister to be involved in program review and decision-making. The benefits from this top-level engagement includes access to resources and fewer bureaucratic obstacles, but drawbacks include political interference and more reporting requirements that impact project management.

- **Trial batch production runs followed by upgrading to improved variants:** A distinguishing pattern of Chinese weapons development programs in recent years has been the manufacturing of very small batches of platforms, often numbering no more than one or two items, which are put into operational service. This is followed with upgraded variants that are also produced in small numbers until the end-users are satisfied and allows for larger production runs.

These features of the accelerated acquisition model can be found in a number of current Chinese weapons development programs. They include the Chengdu J-20 and Shenyang J-15 carrier-borne fighter aircraft, Type 052 DDG, and the Xian Y-20 heavy-lift transport aircraft (see Table 1). The 052 DDG and J-15 fighter programs are prime examples of this fast track model as they have compressed research cycles to Milestone A (4–5 years for the 052 DDG and 2–3 years for the J-15), but lengthy periods for technology, engineering, and demonstration to low rate initial production.

Table 1. Acquisition Cycles for Four Chinese Fighter Aircraft, Transport Aircraft, and Warship Programs

	J20 Fighter	Luyang-Class 052C/D DDG	J-15 Fighter	Y-20 Transport
Preliminary Research to Milestone A	9 Years (1998–2007)	052C: 4–5 Years (1997/8–2001)	2–3 Years (2005–2007/8)	8 Years (2000–2007)
Technology & Engineering Development to Milestone B	9–10 Years (2007–2016/7) Maiden Flight 2011	052C: 7 Years: Initial 2 Years (2001–2003) Followed by Another 5 Years (2005–2010)	9–10 Years (2007/8–2016) Maiden Flight 2009	9 Years (2007–2016) Maiden Flight 2012
Manufacturing Status	MRL 7 LRIP Forecast 2017	052D: MRL 9–10 FRP Begun 2015	MRL 7–8 LRIP Forecast 2020	MRL 7–8 LRIP Begun 2016
Field Deployment	Forecast 2018	052C: 2005 052D: 2014	Pilot Training & Testing since 2015; Operational Deployment 2020	First Aircraft Accepted by PLA Air Force in 2016
Foreign Inputs	Indigenous Platform, Foreign Engines	Indigenous Platform and Armaments, but Heavily Influenced by Russian Designs and Armaments (SAMs)	Reverse Engineered Version of Russian Su-33	Design & Technology Inputs from Ukraine & Russia, especially from IL-76
Upgrading	None Yet	5 Year Gap Between 052C #2 & #3; 052D #1 Followed Immediately After 052C #6	Reports of Electronic Warfare Variant	None Yet
Total Acquisition Period	18–19 Years	052C: 11–12 Years	11–13 Years	17 Years

This accelerated Chinese acquisition process has similarities with the more advanced U.S. defense acquisition system. The United States has also been engaged in efforts to speed up its defense acquisition processes, especially in the past couple of decades in response to its long-term wars and military campaigns in Afghanistan, Iraq, and other parts of the Middle East (DoD, 2015, p. 13). A 2016 study of accelerated U.S. defense acquisition programs by the Institute for Defense Analyses (IDA) identified several types of fast track mechanisms (Van Atta, Kneece, & Lippitz, 2016):

1. **Time-constrained acquisition:** Cost and performance are the primary drivers of acquisition programs, but more attention is now being paid to schedule needs, including to make it a key performance parameter and management priority;
2. **Crash programs:** These are extremely urgent and high priority programs managed by specially created high-level entities that involve the development of early stage technologies with potentially far-reaching disruptive impact. A U.S. example is the F-117A stealth fighter and a Chinese example is the development of its nuclear and ballistic missile programs in the 1950s–1970s;

3. **Rapid acquisition programs:** These projects are aimed to meet urgent operational requirements but involve the development of mature "off-the-shelf" capabilities and are managed by entities located within the regular defense acquisition system. The MRAP is a U.S. example and the DF-21C anti-ship ballistic missile is a Chinese example;
4. **Early fielding experiments:** This refers to promising technologies that are emerging or ready for operational use but lack interest from end-users for acquisition, so the defense acquisition system or defense companies build operational prototypes to experiment. The Global Hawk, Predator, and Reaper unmanned aerial vehicles are U.S. examples, while the Shenyang FC-31 stealth fighter is a Chinese early fielding experiment.
5. **Spiral/evolutionary acquisition:** This is a cyclical or iterative approach that allows for the incremental development of new capabilities, especially for information technology and software projects.

The IDA study identifies top-level leadership support, management, and intervention as the most important factor for the success of U.S. accelerated programs, which is also the case for Chinese programs.

Opportunities and Progress in China's Defense Technological Development and the Role of the Defense Acquisition System

China's defense industry has been enjoying a remarkable renaissance in its fortunes since the turn of the 21st century. Driven by leadership concerns of mounting challenges to the country's external security environment and rapid advances in the global technological order, investment into research, development, and acquisition has soared, greater efforts are being made to acquire and absorb foreign technologies, and the existing defense innovation system is being remade.

This has resulted in significant improvements in technological, economic, and industrial performance. Defense corporations are posting ever-bigger record annual profits, and the armaments research and development pipeline is bulging. The aviation sector, for example, is simultaneously engaged in the development or production of more than half a dozen combat and transport aircraft, while the shipbuilding industry has at least four active nuclear and conventional submarine programs along with research, development, and construction of aircraft carriers, destroyers, and numerous other surface warships. The PLA Navy is estimated to have laid down, launched, or commissioned more than 60 vessels in 2014 and 2015 (ONI, 2015), and commissioned 18 ships in 2016, including 1 Type 052D DDG, 3 Type 054A guided missile frigates and 6 Type 056 corvettes ("Navy Upgrades," 2017; "New 'Carrier Killer,'" 2017). The space industry is also pursuing a highly ambitious across-the-board development, including manned, lunar, anti-satellite, and satellite projects.

A number of key factors have played instrumental roles behind the improving performance of the defense industry. They include high-level leadership support, a clear well-defined long-term vision backed up with detailed development plans and a more capable acquisition system, the growing role of defense corporations, the nurturing of a defense innovation system and overhaul of the research and development apparatus, and efforts to promote the integration of the civilian and defense economies.

Top Leadership Support

High-level and sustained support and guidance from the political and military leadership elites is essential in the Chinese defense industry's ability to carry out innovation activities. Leadership backing and intervention has been vital in addressing entrenched bureaucratic fragmentation, institutional compartmentalization, and chronic project management problems that cause prolonged delays, decision-making paralysis, and cost overruns. Without outside leadership involvement, there would have been a high chance that many achievements of the defense industry would not have happened, especially the turnaround in the defense industry since the end of the 1990s.

The central leadership's direct and continuing involvement and oversight in the operations of the defense industry and of critical projects is essential. This is often done through the establishment of leadership small groups and special committees. The committed involvement of the country's top leaders is also critical, and the defense industry has been fortunate that Xi Jinping has taken a keen and active interest in defense issues. Xi has paid particular attention to the development of China's defense and overall innovation capabilities as demonstrated by his intensive engagement in defense and science and technology matters. Between November 2012 when Xi assumed power and the end of 2016, he took part in 36 publicly reported events related to military and defense science, technology, and industrial issues (see Figure 1). By comparison, Xi's predecessors such as Hu Jintao and Jiang Zemin made far fewer defense-related site visits during the same periods of their rule.

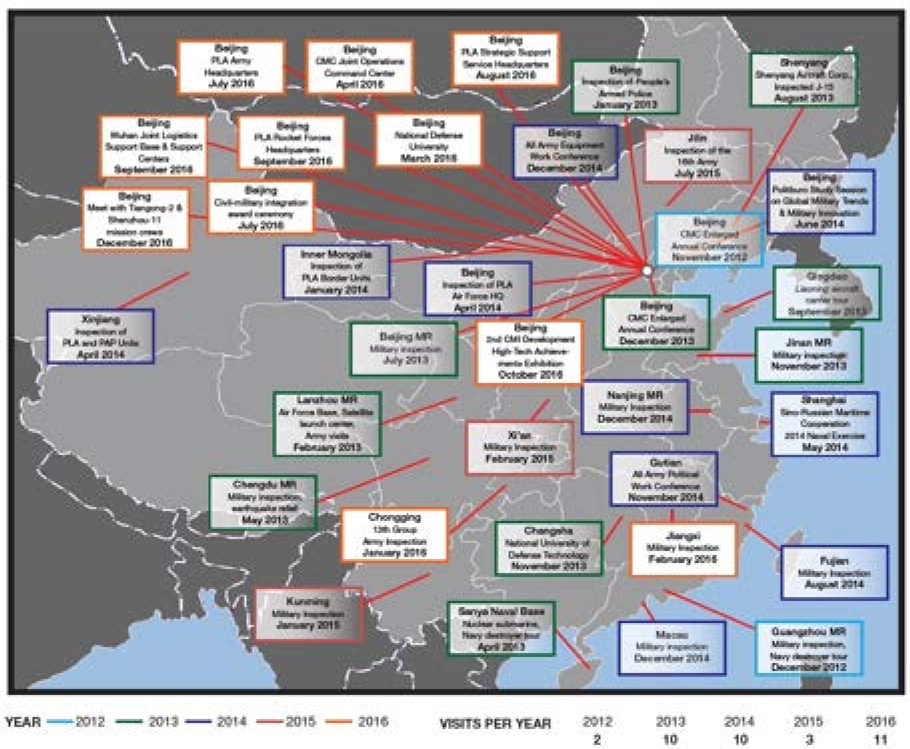


Figure 1. Reported Visits to Military and Defense Science and Technology-Related Facilities Made by Xi Jinping, November 2012–December 2016

Adhering to Medium and Long-Term Strategic Planning

Research and development of China's defense technological capabilities have benefited considerably from the formulation and implementation of medium-term (5 years) and long-term (10–20 years) armament development plans. These strategic and planning guidances provide long-term planning stability, help to mitigate against parochial bureaucratic interests and political intervention, and more rigorously link the weapons acquisition process with threat assessments.

One of the most important of these plans is the 2006–2020 Medium and Long-Term Defense Science and Technology Development Plan (MLDP) that focuses on guiding defense-related basic and applied research and development over a 15-year time horizon (COSTIND, 2006). COSTIND drafted the MLDP in the mid-2000s in coordination with a national medium and long-term science and technology development plan (MLP) and targeted the development of China's defense research and development capabilities and innovation eco-system in a number of areas.

First was the focus on enhancing the capacity for original innovation through the building of defense laboratories, research institutes, and universities to promote basic science along with the development of service support capabilities such as technology transfer and commercialization mechanisms. A second priority was the building of a robust governance regime, especially in areas such as regulations, intellectual property protection, and establishing a comprehensive standards system. A third area targeted was civil-military integration. The overall goal of the MLDP was to significantly narrow the technological gap with the world's leading defense technological powers by 2020.

Another important medium-term defense science and technology development plan is the New High-Technology Program, which is also known as the 995 Program, in reference to the U.S. bombing of the Chinese embassy in Belgrade in May 1999 that was the spark for this project.[4] The Chinese leadership's reaction to the attack was to sharply intensify efforts to develop strategic weapons systems, or what the PLA terms "Assassin's Mace," or *Shashoujian* capabilities. According to Gen. Zhang Wannian, who was a CMC vice chairman during the Belgrade Embassy crisis, the CMC convened an emergency meeting immediately following the bombing, and one of the key decisions made at the meeting was to "accelerate the development of Shashoujian armaments" (Zhang Wannian Writing Team, 2011, p. 416).

Zhang pointed out that then CMC Chairman Jiang Zemin was especially insistent on the need to step up the pace of development of *Shashoujian* mega-projects, saying that "what the enemy is most fearful of, this is what we should be developing" (Zhang Wannian Writing Team, 2011, p. 419). As the "enemy" was the United States, the implication was that the defense and strategic science, technology, and innovation systems should be engaged in developing asymmetric capabilities targeting U.S. vulnerabilities.

The Central Role of Defense Conglomerates

The revival of China's 10 major state-owned defense corporations since the beginning of the 21st century has had a major impact in shifting the center of gravity for research, development, and innovation from research academies and universities towards enterprises. These conglomerates, each of which have between 100 to more than 200 subsidiaries, have sought to transform themselves from loss-making quasi-state

---

[4] There is no official Chinese acknowledgement of the 995 Project, but there are occasional allusions to it in media reports, writings by Chinese military analysts, résumés of Chinese scientists, and project listings of university laboratories and companies engaged in defense-related work. See, for example, Zeng (2009).

bureaucracies to become more market-driven enterprises. They have been slimmed down, allowed to shed heavy debt burdens, and given access to new sources of investment, especially from the capital markets.

Combined with a strong pickup in defense and civilian orders, these companies have become highly profitable since the mid- to late 2000s. Around two-thirds of the defence industry's annual revenue comes from civilian operations, such as automobiles and white goods. The aviation, space/missile, defense electronics, and naval sectors have been the chief beneficiaries from this rising tide of defense procurement, while the ordnance industry has enjoyed considerable success from sales of civilian products such as motor vehicles. These corporations are now engaged in an ambitious expansion strategy to become global arms and strategic technology champions.

Building of a Defense Innovation System and Research and Development Base

The Chinese defense innovation system, and especially its research and development component, has been undergoing a significant overhaul and expansion to meet growing demand for its services from the PLA and also as part of a larger development of the national innovation system. The development of a robust defense R&D system is a top priority in defense science and technology development plans such as the MLDP, which emphasizes a number of key goals. A top priority is the shifting of ownership and funding of key portions of the state-controlled defense R&D apparatus to the country's defense conglomerates. The primary goals of this reform include (1) reducing the dependence of the R&D apparatus to state funding; (2) increasing the amount of investment that firms devote to R&D, especially in applied and commercial development; and (3) speeding up the exploitation and commercialization of proprietary R&D output.

Another high-level priority is the development of an extensive defense laboratory system that would pave the way for long-term technological breakthroughs. Around 90 laboratories belonging to both the defense industry and PLA have so far been established. It will take some time though before these research outfits are able to conduct high quality R&D because they lack experienced and top-rated scientific personnel.

Civil-Military Integration

Intensifying efforts have been made since the early 2000s to forge close linkages between the civilian and defense economies to allow the defense industry to gain access to more advanced and more globalized civilian sectors. This has led to the development of some modest functional and geographical pockets of civil-military activity since the early to mid-2000s. The electronics, information technology, high technology, and automotive sectors have been in the vanguard.

Another area of growing CMI activity is the competitive opening up of the defence research, development, and acquisition system to the private sector that until a few years ago was the exclusive preserve of the 10 state-owned conglomerates that monopolized the aviation, space and missile, ordnance, nuclear, electronics, and shipbuilding sectors that make up the defence industrial base. More than 1,000 private firms have so far received licenses that allow them to bid for contracts, although it is likely that the overwhelming flow of business still goes to the established state giants because they have deep-seated connections in a non-transparent and under-regulated system.

The use of capital markets to fund the development and production of weapons projects is the third area in which CMI initiatives are being pursued, and this has potentially the most significant near and longer term impact on innovation. While defence companies have been allowed to list subsidiaries on stock markets since the 1990s, this was limited to their non-defence operations. This changed in 2013 when SASTIND permitted firms to issue share placements using military assets as securitization.

The high level of state commitment to the defence industry shows few signs of weakening anytime soon, despite the noticeable slowing of growth in the national economy in the past couple of years. For defence research and development, the investment in national science and technology activities is a useful proxy indicator of political support and the trajectory in growth rates in science and technology. China's research and development expenditure in 2015 was Rmb 1.42 trillion ($208 billion), which was 2.07% of Gross Domestic Product (GDP) and a sizeable increase of 8.9% from the 2014 spending of Rmb 1.34 trillion ("China's R&D Spending," 2016). However, it is not known how much was spent on defence-related activities. The Chinese authorities have set a target for science and technology spending to reach 2.5% of GDP by 2020, which would mean even higher growth rates in budget growth for the next few years.

Constraints and Weaknesses in China's Defense Industry and Acquisition System

The principal constraints and weaknesses that the Chinese defense industry faces at present stem from its historical foundations and the uncertain efforts to overcome the corrosive legacy of its difficult history. The institutional and normative foundations and workings of the Chinese defense industry were copied from the former Soviet Union's command defense industry and continue to exert a powerful influence to the present day. The PLA and defense industrial regulatory authorities are seeking to replace this outdated top-down administrative management model with a more competitive and indirect regulatory regime, but there are strong vested interests that do not want to see any major changes.

Monopolies

One of the biggest hurdles that PLA and civilian defense acquisition specialists point out is the defense industry's monopoly structure. Little competition exists to win major weapons systems and defense equipment because each of China's six defense industrial sectors is closed to outside competition and is dominated by a select handful of state-owned defense corporations. Contracts are typically awarded through single sourcing mechanisms to these corporations. Competitive bidding and tendering only takes place for non-combat support equipment, such as logistics supplies.

An effort in 1999 to inject more competition by splitting corporations that monopolized their sectors into two separate entities did little to curb monopolistic practices because these firms focused on different areas of business in their domains and there was little direct rivalry. These powerful defense firms have subsequently sought to reverse this effort at de-monopolization by finding ways to re-merge or collaborate together. In 2008, the aviation industry made the first and so far only successful challenge by consolidating its two post-1999 entities back into a single monopoly structure. There have been occasional reports that the space and shipbuilding sectors might also seek to re-establish a single holding company arrangement.

Bureaucratic Fragmentation

A second serious weakness that has seriously handicapped the effectiveness of the Chinese defense industry is its bureaucratic fragmentation. This is a common characteristic of the Chinese organizational system, but is especially virulent within the large and unwieldy defense sector. A key feature of the Soviet approach to defense industrialization that China imported was a highly divided, segmented, and stratified structure and process. There was strict separation between the defense and civilian sectors as well as between defense contractors and military end-users, compartmentalization between the conventional defense and strategic weapons sectors as well as among the different conventional defense industrial sub-sectors, and division between research and development entities and production units. Key reasons for this excessive compartmentalization include an obsessive desire for secrecy and the powerful influence of the deeply ingrained Chinese model of vertical functional systems (tiao tiao) that encouraged large-scale industries like those in the defense and supporting heavy industrial sectors such as iron and steel and chemicals to become independent fiefdoms.

This severe structural compartmentalization is a major obstacle to the development of innovative and advanced weapons capabilities because it requires consensus-based decision making that is carried out through extensive negotiations, bargaining, and exchanges. This management by committee is cumbersome, risk-adverse, and results in a lack of strong ownership that is critical to ensure that projects are able to succeed the thicket of bureaucratic red tape and cut-throat competition for funding.

When the Chinese authorities in the late 1950s began to pursue the development of strategic weapons programs such as nuclear weapons and ballistic missiles, they recognized that the fragmented nature of the defense industrial economy represented a potentially fatal weakness, so they designed a special high-level leadership arrangement called the Central Special Committee (CSC) to provide the decisive leadership support needed for high-priority strategic projects (Cheung, 2012). The CSC played a central role in ensuring the successful development of China's strategic weapons capabilities, so much so that the Chinese authorities resurrected this leadership group in the late 1980s to oversee the initial development of key strategic programs. The CSC has played an important role in the early development of the Shenzhou manned space project, for example, and has been mentioned in other major strategic technology programs such as nuclear submarine development and other space projects.

This entrenched bureaucratic fragmentation is a prominent feature of the armament management system. Although the GAD was one of the PLA's four general headquarters departments with a seat on the CMC, it was only responsible for managing the armament needs of the ground forces, People's Armed Police, select space programs, and the militia (Mao, 2012, p. 46). The navy, air force, and Second Artillery had their own armament bureaucracies, and competition is fierce for budgetary resources to support projects favored by each of these services. This compartmentalized structure serves to intensify parochial interests and undermines efforts to promote joint undertakings.

The defense acquisition system also suffers from compartmentalization along many segments of the acquisition process. Responsibilities for research and development, testing, procurement, production, and maintenance are in the hands of different units and under-institutionalization has meant that linkages among these entities tend to be ad hoc in nature with major gaps in oversight, reporting, and information sharing (Liu & Wang, 2009). The fragmented nature of the acquisition process may help to explain why Hu Jintao was apparently caught by surprise by the first publicized test flight of the J-20 fighter aircraft that occurred during the visit of U.S. Defense Secretary Robert Gates in January 2011 (Bumiller & Wines, 2011; Pomfret, 2011).

Weak Acquisition Management Mechanisms

A third major weakness is that the PLA continues to rely on outdated administrative tools to manage acquisition projects with defense contractors in the absence of the establishment of an effective contract management system. The PLA did implement the use of contracts on a trial basis in the late 1980s with the introduction of a contract responsibility system (Cheung, 2008, pp. 83–85). These contracts are administrative in nature though and have little legal rights because of a lack of a developed legal framework within the defense industry. Consequently, contracts are vague and do not define contractual obligations or critical performance issues such as quality, pricing, or schedules. Contracts for complex weapons projects can be as short as 1–2 pages, according to analysts.[5]

Moreover, the military acquisition apparatus is woefully backward in many other management approaches and tools that it uses compared to its counterparts in the United States and other advanced military powers. It has yet to adopt total life-cycle management methods, for example, and many internal management information systems are on stand-alone networks that prevent effective communications and coordination. One analyst said that this often meant that the only way for project teams to exchange information was through paper transactions.[6]

Outdated Acquisition Pricing Regime

A fourth serious weakness is the lack of a transparent pricing system for weapons and other military equipment, representing a lack of trust between the PLA and defense industry. The existing armament pricing framework is based on a “cost-plus” model that dates to the planning economy, in which contractors are allowed 5% profit margins on top of actual costs (Mao, 2012, pp. 158–159). There are a number of drawbacks to this model that holds back efficiency and innovation. One is that contractors are incentivized to push up costs as this would also drive up profits. Another problem is that contractors are not rewarded with finding ways to lower costs such as through more streamlined management or more cost-effective designs or manufacturing techniques. Contracts rarely have performance incentives, which discourages risk-taking and adoption of new innovative approaches.

To address this long-standing problem, the PLA, Ministry of Finance, and National Development and Reform Commission held a high-level meeting on armament pricing reform in 2009 that concluded that the outdated pricing system had seriously restricted weapons development and innovation (Zong & Zhao, 2009). A number of reform proposals were put forward: (1) provide incentives to contain costs; (2) switch from accounting procedures that focus on ex post pricing to ex ante controls; and (3) expand from a single pricing methodology to multiple pricing methods.

At the beginning of 2014, the GAD announced that it would conduct and expand upon pilot projects on equipment pricing. These reforms include the strengthening of the pricing verification of purchased goods, improving cost controls, shifting from singular to plural pricing models, from “after-purchase pricing” to “whole process pricing,” and from “individual cost pricing” to “social average cost pricing” (“Armament Work,” 2014). These represent modest steps in the pricing reform process, but the PLA will continue to face fierce opposition from the defense industry on this issue.

Corruption

A fifth impediment is corruption, which appears to have thrived with the defense industry’s uncertain transition from centralized state planning to a more competitive and

---

[5] From an interview with a PLA acquisition specialist, Beijing, November 2011

[6] Ibid.

indirect management model.[7] PLA leaders have highlighted the RDA system as one of a number of high-risk areas in which corruption can flourish along with the selection and promotion of officials, the enrollment of students in PLA-affiliated schools, funds management, and construction work ("PLA Gets Tough," 2014).

At the PLA's annual conference on military discipline inspection work in January 2014, CMC Vice-Chairman General Xu Qiliang, who heads the PLA's anti-corruption efforts, pointed out that armament research, production, and procurement was one of two areas that required "better oversight" ("CMC Vice Chairman Stresses," 2014). The other area that Xu highlighted was construction projects, which has been plagued by a number of high-profile corruption scandals in recent years.

The almost complete absence of public reporting on corruption in the defense industry and acquisition system means that the extent of the problem is not known. Military authorities justify this lack of transparency as many of the cases are likely to involve classified programs. In the latest anti-corruption crackdown that began with Xi Jinping's ascent to power at the 18th Party Congress in November 2012, there have only been a handful of cases of defense industry executives being arrested on corruption charges (e.g., see "Wu Hao, Deputy General Manager," 2014).

The Next Stage in China's Defense Technological Transformation: Formulating New Long-Term Plans and the Reform of the Defense Acquisition System

The Xi Jinping administration signaled its intention to carry out a major overhaul of the defense industry as part of an ambitious national program of economic and military reforms at the Third Plenum of the 18th Party Congress in 2013. A flurry of activity since then by defense industrial decision-makers has produced new medium and long-term defense industrial development strategies, plans, and institutional arrangements that collectively represent a potentially key turning point in the defense industry's evolution from an innovation follower to becoming an original innovation leader. After almost two years of investigation, a reform plan was approved and released at the CMC Working Conference on Reform in November 2015, which marked the formal start of the implementation of the most far-reaching structural reform of the PLA in its history ("Documentary of the Design," 2015).

While these reforms were targeted at the PLA's central, regional, and service commands, it also had important implications for the armament management system, which plays a highly influential role in defense science, technology, and industrial matters. At the end of 2015, the PLA's armament system underwent a far-reaching reorganization ("Central Military Commission," 2016):

- The GAD was reorganized into the CADD and given responsibility for the centralized unified management of the military armament system ("Ministry of National Defense," 2016).

[7] *Corruption* is defined broadly in China as covering the improper behavior of state, party, or military officials, but the more common Western definition is the abuse of public office for personal gain in violation of rules.

- The GAD Science and Technology Committee was elevated to a commission-level rank reporting directly to the CMC and renamed as the CMC Science and Technology Committee (CSTC).

Although it will take some time before these reforms are fully implemented and can be adequately assessed, some initial speculative thoughts can be offered. First, the promotion of the CSTC from the GAD to the CMC demonstrates that the Chinese military authorities, and especially Xi, are increasingly serious about engaging in higher-end STI activities and establishing a high-level coordinating mechanism through the CSTC to provide operational leadership and guidance.

Second, the ability of the new CADD to carry out its mandate of providing centralized management of the armament system looks to have a greater chance of success than the GAD, which was hamstrung by its institutional bias towards the oversight of the ground forces. The nature of the relationship between the CADD and the armament departments belonging to the service arms will be critical in determining how much jointness versus compartmentalization there will in the PLA's armaments development. The authority and influence of the CADD will benefit with the appointment of GAD Director Gen. Zhang Youxia as its new head, who reportedly has close ties with Xi ("Former GAD Director," 2016).

In parallel, the state defense industrial bureaucracy has formulated new strategies and plans for a significant adjustment to the defense industry as well as to chart its medium and long-term transformation. One of these key plans is the 13th Defense Science, Technology, and Industry Five Year Plan (13th Defense S&T FYP). This plan was issued at the beginning of 2016 and sets out six key tasks to 2020: (1) facilitating so-called "leapfrog" development of weapons and military equipment; (2) enhancing innovation capabilities in turnkey areas; (3) improving overall quality and efficiency; (4) optimizing the structure of the defense industry and vigorously promoting civil-military integration; (5) accelerating the export of armaments and military equipment; and (6) supporting national economic and social construction ("2016 National Defense Science," 2016).

Compared to its predecessor, the 13th Defense S&T FYP has a stronger focus on the development of high-technology weaponry and civil-military integration. It also signals a significant shift in the direction of defense industry development from absorption and re-innovation to giving greater emphasis to original innovation. The 13th FYP also shows that China is seeking to build on the inroads it has been steadily making in the international arms market. Chinese arms sales have almost doubled over the past five years, according to the Stockholm International Peace Research Institute and it supplies arms to 37 countries, although three-quarters of the exports were within the Asia-Pacific region, led by Pakistan, Bangladesh, and Myanmar ("China Almost Doubles," 2016).

A long-standing Achilles' heel of the Chinese defense industry being addressed by defense planners is a lack of higher-end manufacturing capability. Currently, SASTIND is in the process of preparing a "2025 Defense Science and Technology Plan" that will align closely with the national-level "Made in China 2025 Advanced Manufacturing Plan" and "Internet Plus" Plan which are aimed at lifting the overall level of the country's industrial equipment manufacturing base and curtailing excessive dependence on foreign core technology and products. The defense industry features prominently in the Made in China 2025 plan, especially the space and aviation sectors ("Defense 2025 Is Coming Soon," 2015).

Implications for the United States

The emergence of the Chinese defense industry and acquisition system as an increasingly capable and peer competitor has enormous implications and challenges for the United States, which can be assessed at three levels: geo-strategic, industrial, and acquisition.

At the geo-strategic domain, the two countries are increasingly engaged in an escalating arms competition with each other. While the Third Offset Strategy is focused at rectifying the overall global erosion in U.S. defense technological pre-eminence, the top challenge over the next 25–30 years comes from the "great powers" of Russia and China. Although the Pentagon is deeply concerned with Russian aggression in the short to medium term, China "embodies a more enduring strategic challenge," according to U.S. Deputy Defense Secretary Work ("Work Outlines Key Steps," 2015).

The Third Offset Strategy has a number of characteristics, in which China looms large as the "pacing threat":

- **Conventional deterrence against great powers:** The central tenet of the U.S. strategy is to develop a dominant conventional deterrent against Russia and China that reduces the chances of major military conflict between them.
- **Asymmetric competition:** Avoid competing in quantitative arms races with potential adversaries and instead focus on developing technologically superior quality that would compensate for the numerical superiority enjoyed by these rivals.
- **Strategy based, technology-oriented:** While technology is important, operational strategies and organizational constructs are also key elements in gaining advantages against numerically stronger opponents.
- **Cost Imposition:** With constrained resources, the United States is looking at ways to shift the cost equation that is heavily in favor of China by seeking to impose higher costs, such as forcing the Chinese to invest in areas that are extremely expensive and in which the United States has a technological edge, like in autonomy.
- **Operational level of war:** The primary focus of the initiatives is in the operational planning and conduct of campaigns that consist of assigning missions, tasks, and resources to military organizations. The principal operational concerns that the Defense Department has are as follows (Martinage, 2014, pp. 23–32):
    1. Growing vulnerability of its global system of military bases, especially those that are close to major potential adversaries in the Asia-Pacific and Europe;
    2. Increasing ability of opponents to detect, track, and engage U.S. aircraft carriers and other major surface warships at extended ranges from their coasts;
    3. Build-up of modern integrated air defense systems that is making it increasingly difficult for U.S. and allied airpower to enter into contested opposition airspace;
    4. Militarization of space that no longer makes it a sanctuary from military conflict.

While the Third Offset Strategy is still at an early stage of development, it does signal that the United States has unambiguously taken its first consequential steps in engaging

China directly in defense technological competition. From a U.S. defense acquisition perspective, this strategy is being operationalized in the Long-Range Research and Development Planning Program, which is modeled on an effort started in the 1970s when the United States successfully offset Soviet military numerical superiority with disruptive technological capabilities such as stealth and precision strike ("DoD Seeks Future Technology," 2014).

While there is little open discussion by Chinese military or civilian officials about the technological threat posed by the United States, they have been responding vigorously at the defense acquisition level since the end of the 1990s, most notably with the 995 Plan, which can be viewed as the Chinese counterpart to the Third Offset Strategy.

At the industrial level, the advances that the Chinese defense industry has accomplished over the past two decades have been impressive, but can they continue at such a rapid pace and in which direction will they lead? If the critical enabling factors that have been instrumental to this progress are still in place, then the prospects look encouraging for China's continued defense technological transformation.

Two particularly key drivers are leadership support and the threat environment. Xi Jinping will almost certainly stay at the leadership helm until the 20th CCP Congress in 2022, so leadership support for the defense industry will remain strong. China's external security environment will remain complicated because of sovereignty disputes and structural competition with the United States and regional neighbors such as Japan. Moreover, the PLA's efforts to build up its long-range power projection capabilities to support its increasingly global ambitions look set to continue. These factors make it likely that the generous levels of funding that the defense industry has received will continue at least over the course of the 13th Five Year Plan to the end of this decade.

However, the continued progress in the development of China's defense technological capabilities rests on troubled foundations. The structural weaknesses of the defense industry makes it at serious risk of falling into a trapped transition, whereby key components are left unreformed or only partially reformed because of strong opposition from powerful interest groups. The negative consequences from this selective reform process has so far been masked by the abundance of resources flowing through the defense industry. But any tightening in budgets because of slowing economic growth could expose the fragilities of this deeply fragmented and flawed system.

At the defense acquisition level, the impact and implications of Chinese developments for the United States primarily revolve around competition in four critical areas: cost, schedule, performance, and innovation. Which country's acquisition system is producing outcomes that are faster, cheaper, better, and bolder than the other side? The Chinese defense acquisition system today is competitive or ahead in cost and schedule, and is behind but narrowing the gap in performance and innovation. As long as the United States is able to maintain a healthy lead of at least one generation or more in the technological capability and innovation of its weapons systems, this offsets China's advantages in schedule and cost.

But if China is able to succeed in narrowing the overall performance capability and innovation gap to within one and even half a generation and continues to maintain a decisive edge in schedule and cost, then it will have the upper hand in the acquisition competition with the United States. A key question is whether the Chinese system is able to be faster and cheaper even as it becomes better and bolder.

A central reason that the Chinese defense industry has been able to keep costs down and accelerate the pace of acquisition is because it has operated on an absorption-based, good-enough development model. But as the Chinese defense industry transitions to more of an original innovation-based, higher end development framework, risks grow significantly and this will impact the costs and pace of the acquisition process. The underdeveloped Chinese defense acquisition system could very likely find itself overwhelmed and lacking the expertise, experience, and organizational, business, and management tools to manage an advanced technology and innovation enterprise. One key exception is a select number of projects that come under special attention and oversight from the highest levels of the civilian and military leaderships. But they are the exception rather than the rule of the Chinese defense acquisition system.

References

2016 National Defense Science, Technology and Industry Working Conference was held in Beijing. (2016, January 9). *State Council.* Retrieved from http://www.gov.cn/xinwen/2016-01/09/content_5031770.htm

Armament work: It is the right time for reform and innovation. (2014, February 13). *Liberation Army Daily.*

Bumiller, E., & Wines, M. (2011, January 12). Chinese Army tests jet as Gates visits. *New York Times*.

Central Military Commission issues "opinions concerning deepening the reform of National Defense and the Armed Forces." (2016, January 1). *Xinhua Domestic Service*.

Cheung, T. M. (2008). *Fortifying China*. Ithaca, NY: Cornell University Press.

Cheung, T. M. (2012, July). *The special one: The Central Special Committee and the structure, process, and leadership of the Chinese defense and strategic dual-use science, technology and industrial triangle*. Unpublished conference paper.

China almost doubles weapons exports over past five years, with Pakistan biggest buyer: Think tank. (2016, February 22). *South China Morning Post*.

China, Russia to co-develop heavy-lift helicopter in 2016. (2015, September 10). *China Daily.*

China's R&D spending rises, still lags behind developed nations. (2016, November 12). *Xinhua News Agency*.

CMC vice chairman stresses effective anti-corruption. (2014, January 17). *Liberation Army Daily.*

Commission of Science, Technology, and Industry for National Defense (COSTIND). (2006, May 29). *Outline of Defense Medium and Long Term Science and Technology Development Plan* (国防科技工业中长期科学和技术发展规划纲要).

"Defense 2025" is coming soon: Aero-engines may become the breakthrough. (2015, June 19). *Xinhua News Agency*. Retrieved from http://news.xinhuanet.com/fortune/2015-06/19/c_127931606.htm

Documentary of the design process of deepening defense and military reform by Xi Jinping and the CMC. (2015, December 30). Retrieved from http://news.sina.com.cn/o/2015-12-30/doc-ifxncyar6047368.shtml

DoD. (2015, January). *Operation of the Defense Acquisition System* (DoD Instruction 5000.02). Washington, DC: Author.

DoD seeks future technology via development plan. (2014, December 3). *DOD News*. Retrieved from http://www.defense.gov/News-Article-View/Article/603745

Dorsett, D. (2011, January 5). Transcript of Defense Writers Group roundtable with Vice Admiral David Dorsett.

Former GAD Director Zhang Youxia becomes new director of CMC Armament Development Department. (2016, January 14). *The Paper*.

Liu, H., & Wang, B. (Eds.). (2009). *National defense scientific research test project management.* Beijing, China: National Defense Industry Press.

Mao, G. (Ed.). (2012). *Introduction to the military armament legal system* (军事装备法律制度概论). Beijing, China: National Defense Industry Press (国防工业出版社).

Martinage, R. (2014). *Towards a new offset strategy*. Washington, DC: Center for Strategic and Budgetary Assessments.

Ministry of National Defense holds news conference on CMC administrative reform and reorganization. (2016, January 11). *China Military Online.*

Navy upgrades missile destroyer. (2017, February 22). *China Military Online.*

New “carrier killer” delivered to fleet. (2017, January 24). *China Daily.*

Ogus, A. (2002, December). *Comparing regulatory systems: Institutions, processes and legal forms in industrialized countries.* University of Manchester: Centre on Regulation and Competition.

PLA gets tough on duty crimes. (2014, December 1). *Xinhua News Agency.*

Pomfret, J. (2011, January 12). Chinese Army tests jet during Gates visit. *Washington Post.*

Saunders, P., & Wiseman, J. (2011). *Buy, build, or steal: China’s quest for advanced military aviation technologies.* Washington, DC: National Defense University, Center for the Study of Chinese Military Affairs.

U.S. Navy Office of Naval Intelligence (ONI). (2015, April). *The PLA Navy: New missions and capabilities for the 21st century*. Washington, DC: Author.

Van Atta, R., Kneece, R. R., Jr., & Lippitz, M. (2016, September*). Assessment of accelerated acquisition of defense programs.* Washington, DC: Institute for Defense Analyses.

Work outlines key steps in Third Offset Tech Development. (2015, December 14). *Defense News.*

Wu, H. (2014, June 4). Deputy general manager of AVIC heavy machinery under investigation for corruption. *Xinjing Bao.*

Zeng, L. (2009, April 30). Investment in defense science and technology. *Science and Technology Daily* (科技日报).

Zhang, Y. (2014, January 9). Speech by Zhang Youxia at General Armament Department Party Committee Enlarged Meeting. *China Defense Industry News,* 9.

Zhang Wannian Writing Team. (2011). *Biography of Zhang Wannian* (张万年传). Beijing, China: Liberation Army Press.

Zong, Z., & Zhao, B. (2009, November 13). Major reform considered in work on the prices of our army’s armaments. *Liberation Army Daily*.

Acknowledgments

This paper is based upon work supported by the Naval Postgraduate School Acquisition Research Program under Grant No. N00244-16-1-0012 awarded by the NAVSUP Fleet Logistics Center San Diego. The views expressed in this paper do not

A Systemic Analysis of Military Equipment Acquisition Among NATO Suppliers: A Proof of Concept Based on a Multi-Layered DSS Approach

Martin Zsifkovits—University of the Federal Armed Forces Munich, Germany [martin.zsifkovits@unibw.de]

Gonzalo Barbeito—University of the Federal Armed Forces Munich, Germany [gonzalo.barbeito@unibw.de]

Dieter Budde—Major General (Ret.), German Armed Forces [dibu11@gmx.de]

Max Krüger—University of Applied Sciences Furtwangen, Germany [max.krueger@hs-furtwangen.de]

Stefan Pickl—University of the Federal Armed Forces Munich, Germany [stefan.pickl@unibw.de]

Abstract

The analysis of military weapon or equipment acquisition is facing numerous challenges, such as rapidly changing and advancing technologies, several cost dimensions—including investment costs, maintenance, or upgrading—as well as public interests. Due to the significant monetary investments required, a transparent and understandable acquisition process is needed. Therefore, especially defense acquisition among NATO partners for the increase of overall efficiency is a desirable goal. We propose a systemic approach, conceptualized under NATO supplying regulations, capable of analyzing the acquisition process of military equipment over multiple layers among NATO partners. The layers differ in their dimensions of analysis. The approach is demonstrated in the paper at hand based on a proof of concept, using artificial data.

Introduction

The NATO Defence Planning Process (2014) describes armaments planning as follows:

> Armaments planning focuses on the development of multinational (but not common-funded) armaments programs. It promotes cost-effective acquisition, cooperative development and production of armaments. It also encourages interoperability, and technological and industrial cooperation among Allies and partners.

Military procurements from allied nations and vice versa are embedded in a complicated network of foreign politics and home affairs, as well as preceding prospective acquisitions' decisions. State armaments procurements are not intended for primary economic purposes. Armaments are given priority in order to fulfill constitutional and political tasks of defense, as well as to safeguard national security and foreign policy interests. Procurement is usually the subject of social and economic policy controversies. Therefore it is necessary to make the underlying considerations and objectives clear within a transparent decision-making process. Future challenges call for cooperation to achieve financial, technical, and/or industrial benefits within the field of acquisition. This requires the possibility of objectively assessing the cooperation in an acquisition. This concept can make complex acquisitions more transparent, effective and more modern. As a result, the capabilities of the armed forces can be further developed through continuous modernization.

In order to meet the requirements of transparent and well informed military acquisition, we introduce the proof of concept for a multi-layered systemic approach that combines qualitative and quantitative methods. As this paper at hand only focuses on proving this concept, we introduce a simplified, artificial acquisition process as use case.

In the last years, acquisition processes became more and more complex and sensitive, as indicated in actual press releases and forum discussions. The report of the NATO-Industry Forum on November 9, 2016, stated: “NATO should further deepen its engagement with Industry through joint concept development, involvement in exercises, experimentation and war gaming, in order to design future solutions together.” Additionally, “The length of the NATO acquisition processes was criticized. One-size-fits-all procurement strategies may no longer (in fact, already does not) address NATO needs adequately” (NATO, 2016). NATO is making efforts to improve the acquisition process. SHAPE Allied Command Operations (ACO) Acquisition Management is an organization that deals with acquisition and is responsible for developing, coordinating, promulgating, and implementing acquisition policies and procedures ACO-wide and may conduct centralized acquisition initiatives. This year for instance, for the procurement of vehicles for the HQ KFOR fleet in Kosovo, Human Resources Data Services (HRDS) support at SHAPE Headquarters and Electronic Warfare Services for ACO Training and Exercises (NATO, 2017). In order to understand and to improve these acquisition processes several description layers should lead to a better understanding. In each layer quantitative and qualitative modeling aspects will be embedded and should be analyzed.

As a first approach a specific normalization procedure might help to generalize such a holistic acquisition process. Our acquisition model is characterized by a multilayered and multistage architecture:

- Multi-layered addresses the different sectors being involved.
- Multistage covers the dynamic behavior of the time-dependent process.

It is the focus of our approach to identify, characterize, and predict the distinguished interfaces. In order to characterize them, we are starting from classical best-practice examples and specify the interfaces more and more.

The resulting management cockpit is characterized by our “Big Five of an Advanced Acquisition Architecture”:

- better coordination,
- better monitoring,
- better interpretation,
- better services, and
- better process-stability.

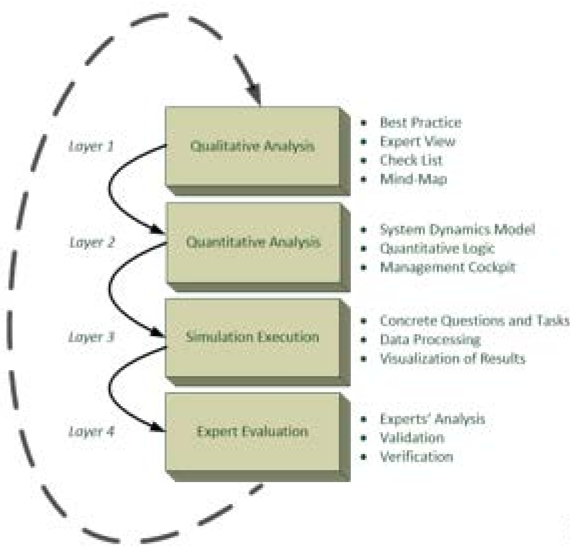
This should be elaborated in the future to develop a comfortable Decision Support System (DSS) as expert analytic framework. The main reason, therefore, is the potential reduction of complexity via such an approach to an acceptable minimum in order to focus on the process itself instead of interpretations of concrete numbers. However, at the end of this article we will discuss the possibility of increasing complexity for real-life evaluations and the automation within a management cockpit as part of a DSS.

This paper is structured as follows: In the next section (Multi-Layered Systemic Acquisition Evaluation: Management Cockpit) we introduce the multi-layered systemic approach for a transparent and understandable acquisition process of drones, comparing several suppliers and their offers on the same scale. The following two sections (Qualitative Analysis [Layer 1] and Quantitative Analysis [Layer2]: System Dynamics) describe the qualitative and quantitative analysis steps of this approach which provide an executable

System Dynamics model (Sterman, 2000). The results of the corresponding simulation are provided in the Simulation Execution (Layer 3) section and discussed in the Validation and Expert Evaluation (Layer 6) section. Possible future work and conclusions are given in the final section.

Multi-Layered Systemic Acquisition Evaluation: Management Cockpit

Due to the high complexity and multiple dimensions of acquisition processes, as discussed above, we propose splitting evaluation of suppliers' bits into several layers. Hereby, every layer represents one step in the evaluation process. We start with a qualitative analysis that brings together experts' opinions with previous best practice examples. This allows us to create a checklist of aspects to be considered in the acquisition evaluation. Furthermore, we propose a visual presentation (e.g., a Mind Map) in order to visualize the interconnections and dependencies between these aspects. This quantitative information is used in a subsequent step for quantitative analysis. Here we suggest the methodology of System Dynamics modelling in order to quantify the previously detected dependencies in the system. For a user-friendly representation of the model and its results, a management-cockpit is suggested. This management cockpit allows the simulation to tackle various research questions at the same time and represents the results adequately. Before taking the final decisions, the qualitative analysis experts from Layer 1 judge the model's output for validation and verification of the results. A graphical representation of this proposed approach is shown in Figure 1.



¶

Figure 1. Complex DSS-System Analysis Over Multiple Layers

These four layers form the main part of the Multi-Layered Systemic Acquisition evaluation and are explained in the following sections in more detail. For demonstration purposes, we provide and use a fictive example of a weapon system acquisition.

Let us consider the acquisition of a novel drone fleet as running example. There might be 10 different suppliers on the national and international market offering such assets. Thereby, prices, technologies, maintenance, and possibilities for upgrades differ among all suppliers. A table of the most important key-parameters for the existing suppliers on the market is shown in Table 1.

Table 1. Key-Parameters of Suppliers' Tender Offers

SUPPLIER	TECHNICAL ATTRIBUTES				MAINTENANCE ATTRIBUTES			SUPPLY SIDE ATTRIBUTES			Cost Upgrade	Price/U
	Equip	Range	Versatility	Load	length	cost/rev	r	n lines	time/unit	stock T(0)		
1	3	6	8	3	20	€ 50.000	2	2	5	10	€ 150.000	€ 1.000.000
2	6	3	7	7	25	€ 70.000	3	4	6	5	€ 50.000	€ 1.200.000
3	8	7	8	5	30	€ 100.000	2	2	4	8	€ 200.000	€ 1.300.000
4	4	4	5	7	10	€ 50.000	2	5	7	5	€ 150.000	€ 800.000
5	4	6	8	7	20	€ 150.000	1	10	12	10	€ 60.000	€ 1.100.000
6	5	5	10	1	30	€ 220.000	1	8	3	0	€ 130.000	€ 900.000
7	4	4	3	6	20	€ 75.000	2	6	5	2	€ 200.000	€ 700.000
8	2	7	4	4	15	€ 75.000	3	7	4	15	€ 80.000	€ 800.000
9	8	8	6	8	10	€ 100.000	3	3	7	4	€ 320.000	€ 2.200.000
10	5	4	6	4	20	€ 80.000	2	3	6	7	€ 120.000	€ 1.100.000

Here, the technological parameters are normalized values on a scale from 0 to 10, with 10 representing the best value. We compare each drone system with respect to equipment, maximum range per flight, versatility, and maximum payload. Furthermore, the average duration per maintenance, the number of maintenances needed per year, and the average cost per maintenance are listed. Suppliers' offers also include information on how many drones are in stock and how many can be built in a given time period and in how many production lines. Even though we are using a simplified scenario with fictive numbers in order to prove our concept, one can see the high degree of complexity underlying to this acquisition decision problem.

Using the systemic multi-layer approach, we demonstrate, subsequently, how the acquisition process can be processed in a transparent way. This should lead to adequate decision support. Therefore, we will use different ways of analysis for a deeper understanding of every offer.

Qualitative Analysis (Layer 1)

Let us assume that the intended useful life of the new drone fleet is 20 years. In the first layer of analysis, decision makers should evaluate best practice acquisitions from the past. This helps to find important aspects in prior cases and allows for learning from prior mistakes. Additionally, experts should be contacted in order to identify a successful acquisition's key issues of assets of this or similar type. The evaluation of these experts' views and best practice examples results go into a checklist that contains relevant factors, such as technological parameters, performance indicators, maintenance circles, upgrade possibilities, or delivery dates. In addition, the checklist should include relevant constraints for these factors, drawn from the call for tenders, for example. In more detail, the checklist may specify upper or lower bounds (e.g., for prices). This allows for reducing the quantity of offers in a first step. With respect to our running example, we assume the following checklist with desired attributes and constraints:

- Minimum level of equipment, range, and load is 3.
- Minimum level of versatility is 5.
- Annual maintenance costs should not exceed €210,000.
- Four drones are needed immediately.
- Acquisition costs should not exceed €1,300,000.

Applying this checklist to the existing offers in Table 1, the number of valid offers substantially reduces. Table 2 shows the suppliers' tender offers with every value outside defined boundaries marked in red. These offers are excluded from further consideration.

Table 2. Reduction of Offers Based on Checklist's Application

SUPPLIER	TECHNICAL ATTRIBUTES				MAINTENANCE ATTRIBUTES			SUPPLY SIDE ATTRIBUTES			Cost Upgrade	Price/U
	Equip	Range	Versatility	Load	length	cost/rev	r	n lines	time/unit	stock T(0)		
1	3	6	8	3	20	€ 50.000	2	2	5	10	€ 150.000	€ 1.000.000
2	6	3	7	7	25	€ 70.000	3	4	6	5	€ 50.000	€ 1.200.000
3	8	7	8	5	30	€ 100.000	2	2	4	8	€ 200.000	€ 1.300.000
4	4	4	5	7	10	€ 50.000	2	5	7	5	€ 150.000	€ 800.000
5	4	6	8	7	20	€ 150.000	1	10	12	10	€ 60.000	€ 1.100.000
~~6~~	5	5	10	1	30	€ 220.000	1	8	3	0	€ 130.000	€ 900.000
~~7~~	4	4	3	6	20	€ 75.000	2	6	5	2	€ 200.000	€ 700.000
~~8~~	2	7	4	4	15	€ 75.000	3	7	4	15	€ 80.000	€ 800.000
~~9~~	8	8	6	8	10	€ 100.000	3	3	7	4	€ 320.000	€ 1.200.000
~~10~~	5	4	6	4	20	€ 80.000	2	3	6	7	€ 120.000	€ 2.200.000

Application of this first analysis step resulted in a reduction of relevant offers by 50%. However, the remaining offers and their underlying information need further analysis. In the following step we outline the individual influences previously evaluated in a graphical representation (e.g., a mind map). This illustration helps to understand interdependencies among various factors and might already provide a qualitative idea of preferences. For our running example in drone acquisition, this presentation is given in Figure 2.

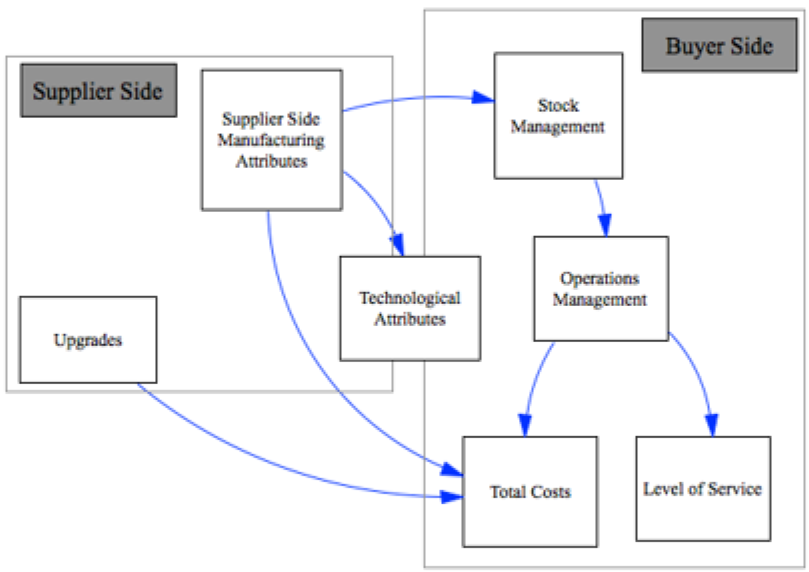


Figure 2. Graphical Representation of Relevant Acquisition Factors

This graphical representation sketches a first qualitative impression of various values' relations, such as monetary costs, service levels, technological attributes, or waiting time according to the stock management. All these factors are of importance for the acquisition and therefore demand for further analysis. In order to make all these values more concrete and comparable among each other, in the next section we transfer the mind map into a quantitative model.

Quantitative Analysis (Layer 2): System Dynamics

Many methodologies for quantitative analysis exist. We decided to apply System Dynamics, which is a computer-aided approach to policy analysis, design, and the definition of corporate strategies. System Dynamics (SD) was developed by Jay W. Forrester at the Sloan School of Management of the Massachusetts Institute of Technology in the 1950s (System Dynamics Society, 2017).

System Dynamics is an established methodology, which has been applied in various areas before. Coyle (1996) provides a practical oriented introduction in qualitative and quantitative modelling with Systems Dynamics. For example in military context, Minami and Madnick (2009) apply System Dynamics for the analysis of combat vehicle accidents in Afghanistan. Fan, Fan, and Chang (2010) provide a modeling of military weapon maintenance supply systems for the analysis of the bullwhip effect in military supply chains, and in Adamides, Stamboulis, and Varelis (2004) System Dynamics is used to support a procurement decision of military aircraft engines.

Considering the given information from our running example and the designed mind map, we construct a System Dynamics model for numerical analysis. In a nutshell, System Dynamics is a modeling and simulation technique for analyzing the behavior of systems over time. Plenty of sources exist with detailed descriptions and information on the methodology, and with a broad audience considered, (e.g., Sterman, 2000; Meadows & Wright, 2008; Pruyt, 2008). All System Dynamics models are formed of three major entities, as described in Zsifkovits et al. (2016):

- **Stocks or levels:** These elements are the foundation of every system. Their main behavior is to accumulate or deplete over time. They can act as delays, lags, buffers, ballast, and sources of momentum in a system. Furthermore, these entities allow inflows and outflows to be decoupled and independent and temporarily out of balance with each other.
- **Material and information flows:** Material flows are the elements that modify the value of a stock: inflows add to the stock and outflows subtract from it. Despite their name, they don't necessarily need to carry physical elements. For example, information flows are a representation of who has access to determined information on the system.
- **Delays:** These entities represent the lengths of time relative to the rates of system changes. These delays may be inherent to the system structure, as the rate on which a stock changes, acting as buffer of the system, or explicitly added, by introducing a time delay entity in the model. (Zsifkovits et al., 2016)

These simple elements all can be composed in such a way that complex systems are easily represented.

Even though that System Dynamics is well suited for qualitative behavioral analysis of systems and policies over time, System Dynamics is not to be used as method for forecasting particular future events, but to provide decision makers and experts a better intuitive feel for improving judgment on the factors influencing success (Forrester, 1961; Forrester, 2007). This is also what we are aiming for in this paper.

For the example at hand, the overall structure of the proposed model is shown in Figure 3. The structure is divided into sub-models in order to make the model more clear and easy to understand. It demonstrates all observed interactions on basis of eight different sub-models and their interactions, separately for each supplier. The sub-models represent the acquisition costs, the total costs (including acquisition, maintenance, and upgrading), technological aspects, stock management, maintenance, upgrades, supplier side conditions, and the level of service. The sub-models are interacting at various points, which is represented by variable names in grey (in Figure 3), which means that they are used in more than one sub-model at the same time. The individual sub-models are subsequently described in more detail.

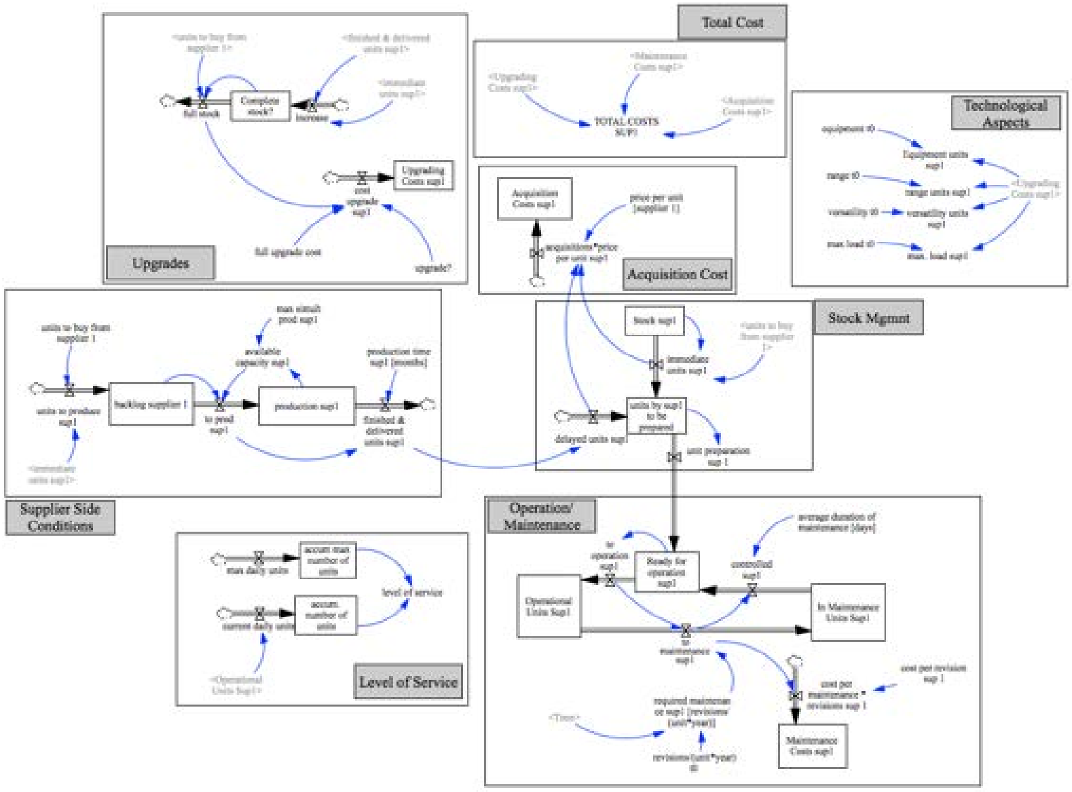


Figure 3. System Dynamics Model

Sub-Model: Acquisition Costs

The first sub-model is very small and counts the overall acquisition costs by simply multiplying the price per unit by the number of ordered units (summing up the immediately delivered and the ones that need to be produced as explained in the sub-model “Supplier Side Conditions”). The interaction of the acquisition costs with the stock management and the maintenance is shown in Figure 4.

Sub-Model: Supplier Side Conditions

As soon as a special quantity is ordered, this model checks how many quantities are available in stock and how many have to be produced. Thereby, the model considers the production time per unit and the existing production lines per supplier. This shows finally, when the quantities in production are available.

Sub-Model: Stock Management

This sub-model combines the existing units in stock and the supplier side conditions in order to see in which time steps which quantities can be delivered.

Sub-Model: Maintenance

The maintenance sub-model considers the different annual service intervals, their individual duration in days, and the costs per maintenance. Every supplier has different maintenance circles, durations of maintenance, and different costs caused by maintenance. Additionally we assume that after 10 years, every product needs an additional maintenance per year. The sub-model as well as its connection to the previously mentioned sub-models is shown in Figure 4.

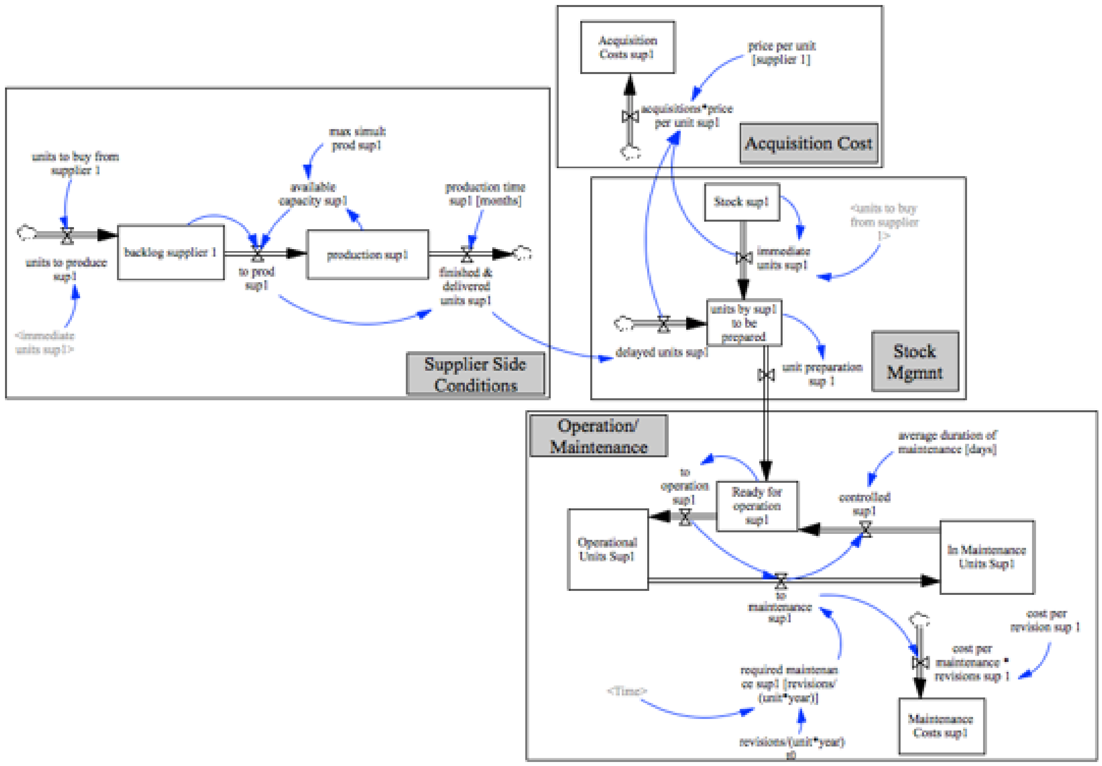


Figure 4. Core Sub-Models

Sub-Model: Total Cost

In this sub-model the total costs are summed up, resulting from acquisition, maintenance, and technological upgrading. Each supplier offers supplementary upgrading, which is considered in future. This upgrading would increase technological parameter and is at cost (as also shown in Table 1). This sub-model is shown in Figure 5.

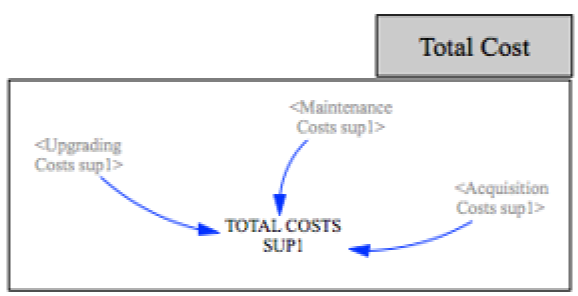


Figure 5. Sub-Model Total Costs

Sub-Model: Technological Aspects

Every product is facing several technological parameters: equipment, range, versatility, and maximal loading capacity. Different suppliers have different views and therefore their products differ in those ranges. For a comparison of their capabilities we normalized the technological values on a scale of 0 to 10. These levels can be improved by technological upgrading at cost in the future, which is explained in the sub-model "upgrading." Initial values are used from the input sheet in Table 1. The sub-model is illustrated in Figure 6.

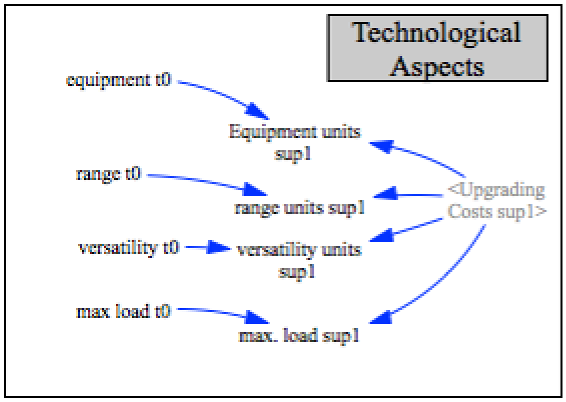


Figure 6. Sub-Model Technological Aspects

Sub-Model: Upgrades

When upgrading the technology of a product, this improves the technological capability by a special amount. To simplify things we assume in this proof of concept, that each technological parameter (equipment, range, versatility, and maximal loading capacity) can always be improved by two units, but keeping the upper bound of 10 units. For a more complex analysis of real cases this would of course differ. However, just for proving the concept suggested in this article, the simplification seems to be adequate. We assume that the upgrades of the fleet are executed as soon as all items are delivered by the supplier. This sub-model is illustrated in Figure 7.

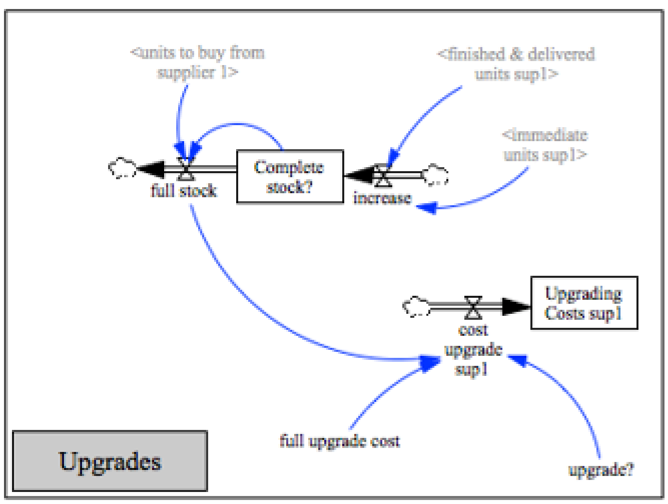


Figure 7. Sub-Model Upgrades

Simulation Execution (Layer 3)

For the remaining five suppliers, which met the constraints in our example, we executed simulation runs and evaluated the results in comparison amongst each other. This should help the decision maker to rank the options for a purchasing decision. We execute the simulations over 20 years in order to project the impact of the acquisition over the full-time horizon of the intended system's use. For the analysis, we tackle three main questions of interest:

- What is the technological status and potential of each product?
- What is the overall cost of acquisition over 20 years, including all relevant factors?
- What is the service level of each supplier over the planning time horizon?

In order to help the decision maker answer these questions, the following results were performed and visualized for the planning horizon of 20 years. However, as different suppliers deliver the orders (at least partly) at a later stage, an analysis over 20 years after delivery of the last piece might also make sense and could be executed.

Decision's Dimension: Technology

The technological parameters in our example are not parameterized dynamically over time. Technological improvement is possible in the same volume over all parameters and suppliers. We introduced this simplification in order not to make this proof of concept too complex. However, results might become far more extensive for real-life scenario analysis if this simplification is omitted. In the scenario at hand we see the initial values and their potential equal improvement over time and parameters in Figure 8.

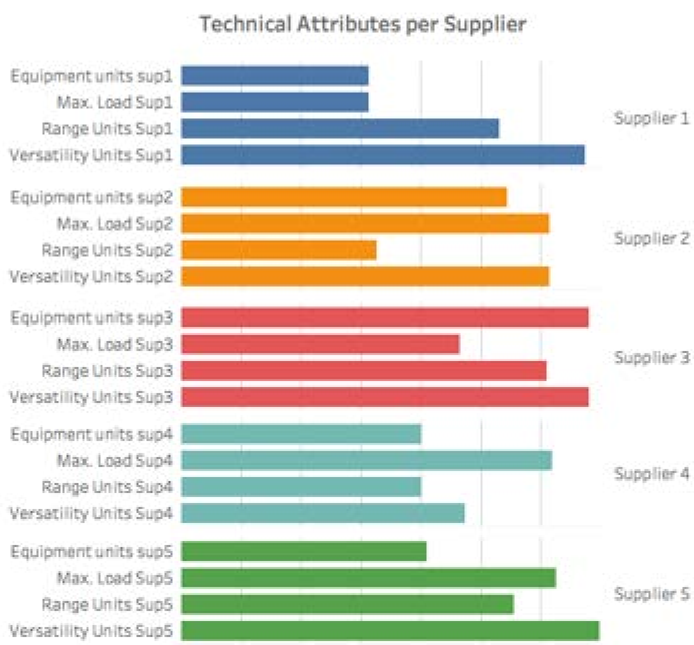


Figure 8. Technological Attributes per Supplier

Decision's Dimension: Costs

Figure 9 shows the results of simulated cost evolution over several dimensions. On top, the overall costs per supplier are accumulated over a 20-year time horizon. It is shown very clearly that the overall acquisition costs vary in a range from about €120 million to over €200 million for the different suppliers. Suppliers 1 and 4 are amongst the cheapest overall, while acquisition from supplier 3 would lead to the highest overall costs. A surprising aspect of this analysis is that suppliers 1 and 4 are amongst the most expensive suppliers in terms of upgrading. However, as these are costs that occur only once, and as these costs are comparably low, this does not influence the overall costs negatively compared to the other suppliers. In the analysis at hand we assume that costs incur after delivery. This leads to different rises of costs over time. First of all, this is essential in order to see when the products are delivered, secondly it might become relevant when discounting the interest to t=0. When having a detailed look at the cost curves, one can see a slight turn upwards at the overall costs, but especially for the maintenance costs. This is caused by the additional annual maintenance needed after 10 years. Several mid- and long-term effects can be analyzed using this systemic approach. Especially with an increase in complexity, these dynamics become increasingly important. The time point when all 40 items are delivered can be easily seen in the costs of upgrading, as all drones are upgraded as soon as they are all delivered from each supplier. This also explains why upgrade costs are a rather sharp increase and therefore leads to the increase in total costs for all suppliers.

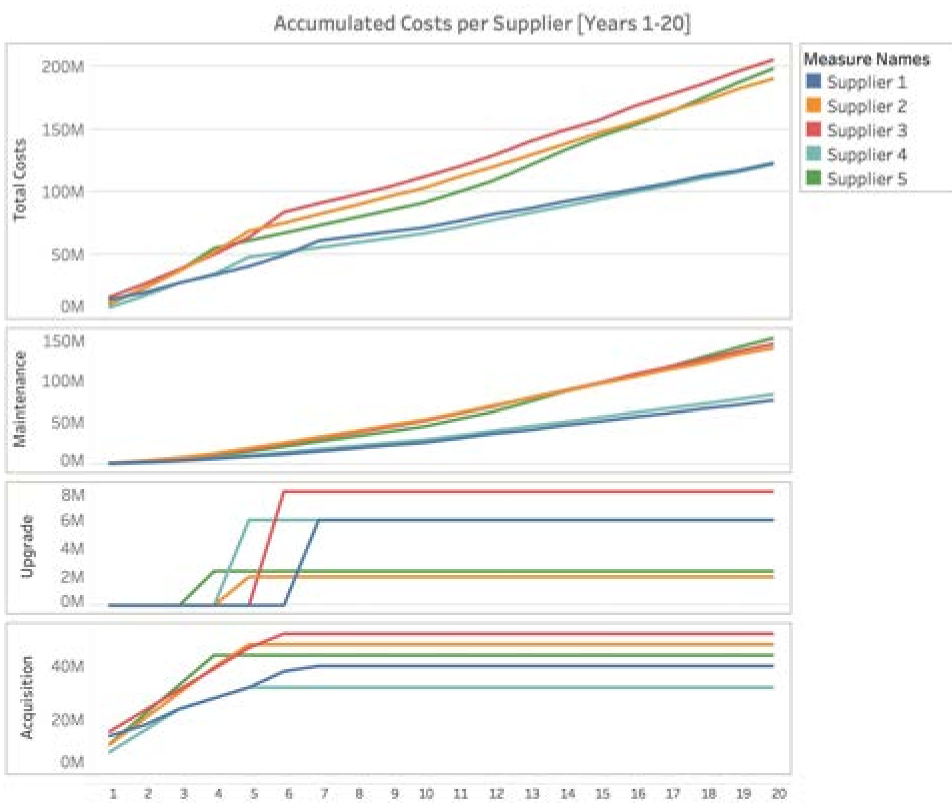


Figure 9. Accumulated Costs per Supplier

Decision's Dimension: Service Level

Another analysis aspect is the degree of expected service by each provider. We define the service being higher if all desired drones can be delivered immediately at t=0. Furthermore, we define a higher degree of service when a system does not have to undergo maintenance too often, the maintenance is shorter, or the unavailability of drones is low. Thus, receiving all 40 drones at t=0 and having no maintenance over the planning horizon, the service level would be 1. We introduce this measure in order to have more information on the offer than only monetary costs and technological parameters. In this case, the highest degree of overall service can be expected by suppliers 4 and 5. While supplier 5, holding the highest degree of service, is amongst the most expensive providers, supplier 4 is the second best in service and is amongst the cheapest suppliers. The service levels of the rather expensive suppliers 2 and 3 are amongst the lowest. The results of the analysis for service levels can be seen in detail in Figure 10.

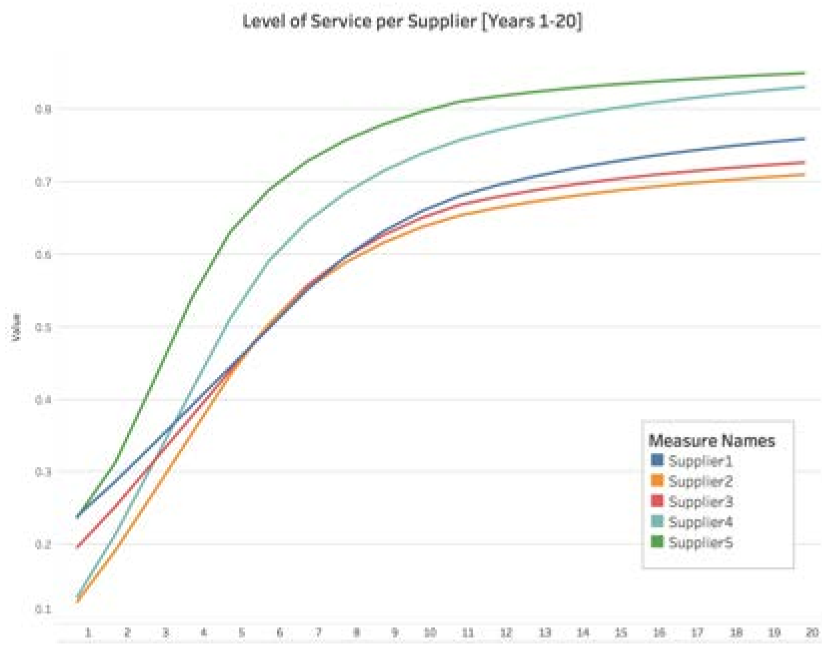


Figure 10. Level of Service per Supplier

The presented model with its eight sub-models was validated and verified in a multiple step approach, combining partial (sub-model) tests and full tests, fixed number execution, and scenario evaluation. At a certain level of abstraction, the model with all its sub-models can be seen as valid for the fictive running example and can be easily modified or extended. This would allow executing simulations for real-life acquisitions beyond this proof of concept.

Validation and Expert Evaluation (Layer 6)

The results of the model executions are finally presented to experts. From a modeler's or analyst's point of view, it is important not to evaluate the results, as this should be done unbiased by the subject matter expert, not by the analyst. The expert is now able to compare all the supplied products over various dimensions. Here it must be noted, that it is not possible to include all possible dimensions in the quantitative analysis. Thus, the decision maker needs to combine quantitative results with other dimensions, such as public interest.

For validation and expert's evaluation of our example, the objective analysis of acquisition allows rationality of procurement programs and thus joint procurement. It is about equipment, weapon systems, and services. The clear definition of common service time and cost parameters allows an economical procurement of armament. The different costs, such as investment, introduction, maintenance, up-grading, and new procurement, as well as technological and organizational support for an acquisition can be better estimated. In addition, cost increases and possible additional costs can be recognized at an early stage. The frequent "low demand for high-quality weapon systems" can be tackled this way.

Investments can be reduced by reviewing existing policies, structures, and procedures for armaments cooperation and acquisition. Furthermore, it can support managing a reduction of forces and a rationalization of military infrastructure or a build-up of new forces and the required equipment, thereby supporting a transformation of forces.

The political, social, and economic interests of states can be bundled through joint acquisition. Thus, synchronizing and utilizing the relevant national line organizations that are in charge of the procurement processes. Using common acquisition deep-rooted national traditions, bureaucratic inertia, diverging industrial interests, and different procurement philosophies might be overcome. This model supports a possible effort for armament cooperation and standardization in bilateral and multilateral bodies and a broad range of armaments programs. This paves the way for technological security, industrial consolidation, and the harmonization of military requirements for systems to be procured. Thus, world-marketable military products can be procured. Joint acquisition of equipment can ensure the required military capabilities, security of supply, cooperative research, and development. Furthermore, unified requirements and capabilities enable cooperation and help to achieve commonality in equipment, support, and usage.

The application of the model can lead to better coordination of procurement procedures for defense and security and better monitoring of supply and service contracts. In particular, a close professional collaboration between the client and the contractor is crucial for success, especially in large projects. The aim of a subsequent phase of realization is to provide the armed forces with timely, ready-to-use products and services for their operations.

This model is used to assess the entire life-cycle of ordnance from the initial stage of research, to propose solutions for their realization and use of the control system, and to separate and exploit them according to rational processes. The different layers of the model, Qualitative Analysis, Quantitative Analysis, Simulation Execution, and Expert Evaluation, and their interdependencies enable an expert to survey the background of the preparation of an acquisition. There are several different areas in which the analysis can be influenced by the provision of expertise. This is particularly true in the context of layer 4. The qualified view of experts in layer 1 is of particular importance, where different experts are required to present their ideas. However, the different perceptions of experts should then be harmonized within the framework of setting key parameters.

Through the model, the decision maker gains the opportunity to get an overview of the entire process of acquisition. Through the objectivity of the processes, the possibility of comparing the parameters and the alternatives, the decision maker finds the basis for a decision. It is possible for him or her to estimate costs and risks and to analyze vulnerabilities. It also ensures early planning security and the possibility of early coordination of further structural and personal measures as well as for training. Overall, the effectiveness and efficiency of the procurement of materials and services, and thus of the armed forces, can be increased by means of the model.

Conclusion and Future Work

The analysis of previous armament projects and armament acquisitions has shown that an improvement in armament acquisition management is required in national and international projects. Acquisition and management of armament projects calls for a culture of leadership in which transparency, integrity, initiative, and responsibility of all project participants are required. With the approach provided in this paper, possible deficiencies can be overcome. The framework conditions for future equipment and procurement to ensure necessary military capabilities and security of supply can be improved. Cooperation efforts and joint defense efforts become more transparent. Common projects such as "Smart Defense" and "Pooling and Sharing" provide the necessary force resources for states who can no longer afford the independent military capabilities through appropriate acquisition. Using the model, possible national selfishness and unwillingness to make necessary investments in the future military capability can be reduced through shared knowledge, parameters for acquisition, and common procurement. All these aspects are candidates to be included in a possible refinement of the proposed acquisition decision support model.

In the current state of this research we introduced only a proof of concept for the systemic analysis of military acquisitions based on a multi-layered decision support system (DSS) approach. In doing so, we introduced an artificial example with reduced complexity. In a further step, we propose to increase complexity in applying the concept to real cases and analyzing real data from historic (or current) acquisition processes. The increase in complexity would also increase dynamics in the simulation and demonstrate the importance of our approach even more. The reason is that several medium- and long-term effects might differ strongly over several suppliers.

A further step in this research might be the concentration and standardization of several groups of cases. For example, weapon acquisitions, vehicle acquisitions, drone acquisitions, or helicopter acquisitions seem to have very individual characteristics, but seem rather homogeneous within these groups. Several DSS for each group of acquisitions might increase the acceptance of our solution for practitioners as it improves the applicability in daily work.

Furthermore, this standardization of the DSS would allow bundling the whole process in a management cockpit. The greatest benefit is thereby the automation of several steps such as the modeling or representation and visualization of results. The management cockpit should be user-friendly software that allows for testing various scenarios or settings within a group of acquisition projects. It should guide the user to an adequate analysis using predefined and dynamic variable inputs. In general, a fully working DSS prototype over all layers is characterized and proposed for future research.

References

Adamides, E. D., Stamboulis, Y. A., & Varelis A. G. (2004). Model-based assessment of military aircraft engine maintenance systems. The Journal of the Operational Research Society, 55(9). 957–967.

Coyle, R. G. (1996). *System dynamics modelling—A practical approach.* Boca Raton, FL: Chapman & Hall/CRC.

Fan, C. -Y., Fan, P. -S., & Chang, P. -C. (2010). A system dynamics modeling approach for a military weapon maintenance supply system. *International Journal of Production Economics, 128*(2). 457–469.

Forrester, J. W. (1961). *Industrial dynamics*. Cambridge, MA: MIT Press.

Forrester, J. W. (2007, June). System dynamics—The next fifty years. *System Dynamics Review, 23*(2–3). 359–370.

Meadows, D. H., & Wright, D. (2008). *Thinking in systems: A primer*. White River Junction, VT: Chelsea Green.

Minami, N. A., & Madnick, S. (2009). Dynamic analysis of combat vehicle accidents. *System Dynamics Review, 25*(2), 79–100.

NATO. (2014). *The NATO Defence Planning Process*. Retrieved from http://www.nato.int/cps/en/natohq/topics_49202.htm

NATO. (2016). *NATO-Industry Forum report*. Retrieved from https://diweb.hq.nato.int/indrel/Shared%20Documents/NIF16_Report_final.pdf

NATO. (2017): *Supreme Headquarters Allied Powers Europe—Acquisition management*. Retrieved from http://www.shape.nato.int/financeandacquisition

Pruyt, E. (2013) *Small System dynamics models for big issues: Triple jump towards real-world complexity*. Delft, Netherlands: TU Delft Library.

Sterman, J. D. (2000) *Business dynamics: Systems thinking and modeling for a complex world* (Vol. 19). Boston, MA: Irwin/McGraw-Hill.

System Dynamics Society. (2017). *U.S. Department of Energy's introduction to system dynamics: Origin of system dynamics*. Retrieved from http://www.systemdynamics.org/DL-IntroSysDyn/start.htm

Zsifkovits, M., Barbeito, G., Pickl, S., & Pauli, G. (2016, July). System dynamics modeling for analyzing business model innovation. In *Proceedings of the First International Conference on Business Management of Technology*. Kumamoto, Japan.

THIS PAGE INTENTIONALLY LEFT BLANK

Additional Papers

On Data Capabilities for Acquisition Management

Richard Wang, MIT and University of Arkansas at Little Rock
Ningning Wu, University of Arkansas at Little Rock

Determining New System Design Requirements to Optimize Fleet Level Metrics Under Uncertainty

Satadru Roy, PhD, Purdue University
Navindran Davendralingam, PhD, Purdue University
William A. Crossley, Purdue University
Parithi Govindaraju, PhD, Purdue University

Realistic Acquisition Schedule Estimates: A Follow-On Inquiry

Raymond Franck, Brig Gen, USAF (Ret.)—Naval Postgraduate School (Ret.)
Gregory Hildebrandt
Charles Pickar—Naval Postgraduate School
Bernard Udis—University of Colorado, Boulder

On Data Capabilities for Acquisition Management

Richard Wang—is Director of the MIT Chief Data Officer and Information Quality Program. He is also the Executive Director of the Institute for Chief Data Officers (iCDO) and Professor at the University of Arkansas at Little Rock. From 2009–2011, Wang served as the Deputy Chief Data Officer and Chief Data Quality Officer of the U.S. Army. He received his PhD in information Technology from the MIT Sloan School of Management in 1985. [rwang@mit.edu]

Ningning Wu—is Professor of Information Science at the University of Arkansas at Little Rock. She received a BS and an MS degree in Electrical Engineering from the University of Science and Technology of China and a PhD in Information Technology from George Mason University. Dr. Wu's research interests are data mining, network and information security, and information quality. She holds certificates of the IAIDQ Information Quality Certified Professional (IQCP) and the SANS GIAC Security Essentials Certified Professional. [nxwu@ualr.edu]

Abstract

Military logistics are responsible for sourcing and providing nearly every consumable item used by military forces worldwide. The process is highly complex; any misplaced decisions have serious cost and security consequences. Central to the entire process is the quality of the data used to make these acquisition decisions. We explore an enterprise approach to improving data capabilities for acquisition management, building upon a cumulative body of knowledge from Chief Data Officer (CDO) and information quality research and practice.

Overview

The success of the U.S. military acquisition process depends in large part on the ability to make data-driven decisions across the entirety of the organization in an efficient and effective manner. By connecting internal management from all branches of the Armed Forces to the Office of the Secretary of Defense (OSD) through improving data capabilities for acquisition management, the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD[AT&L]) provides data stewardship, data access, and data analysis to help improve acquisition insight, management, policies, and processes for trillions of dollars in budgetary assets. These processes are then more accurately measured in reports delivered to the OSD and Congress. Department-wide acquisition includes aspects of performance improvement, budget planning, industry reviews, program milestone decisions, program portfolio reviews, program insight, and portfolio oversight. These deliverables make up the bulk of reporting information and are the responsibility of the USD(AT&L). The mission of improving data capabilities for acquisition management, supported by the Acquisition Resources and Analysis branch of Enterprise Information at the Department of Defense (DoD; see Appendix A), is to provide leadership with timely access to accurate, authoritative, and reliable data supporting acquisition oversight, analysis, and decision-making.

A closer look at the evolution of acquisition policy (Appendix B) at a successful defense acquisition program, as defined by the Office of the Secretary of Defense, shows the definition to be "a program that satisfies national security objectives, provides a balanced force structure, and does not attract undue congressional scrutiny" (Brown, 2010). For the program manager, success also means overseeing a system that is delivered on time, within cost, and meets requirements of their staff. The Quadrennial Defense Report identified four major problems in the DoD's ability to acquire military capabilities in a timely and affordable manner (Brown, 2010):

- Requirements for new systems too often reflect the far limits of current technology, and requirements that continue to increase throughout a program's life cycle.
- The acquisition workforce lacks the trained personnel in the areas of cost estimators, systems engineers, and acquisition managers. This causes problems in the conduct of effective oversight.
- The acquisition process too often encourages overly optimistic cost estimates. Underestimating cost is likely to result in too many programs chasing too few dollars, and cost threshold breaches requiring program terminations and increased reporting to Congress.
- Improvements are needed in the effective and efficient delivery of logistical support to the fighting forces in the field.

These problems outline the heart of Augustine's Laws as they relate to acquisition in the U.S. military. Simply put, there is a cyclical relationship between the acquisition community and contractors that seemingly cannot be broken in the status quo, as the free market pressures that typically would step in and self-regulate supply and demand fluctuations do not exist in the same manner in this closed environment. The budgetary and acquisition problems faced are in desperate need of resolution. The emerging "big data" solutions seem to begin to address pitfalls of military acquisition theory and practice identified by Augustine (Appendix C).

In order to more fully understand the problems facing the growth and realization of improving data capabilities for acquisition management, we must look at the difficult task facing individual program managers, as they try to create the maximum amount of value for their individual program, while consistently facing scrutiny from a variety of sources regarding their respective cost and output levels. Figure 1 shows us the complicated environment of the program manager (Brown, 2010).

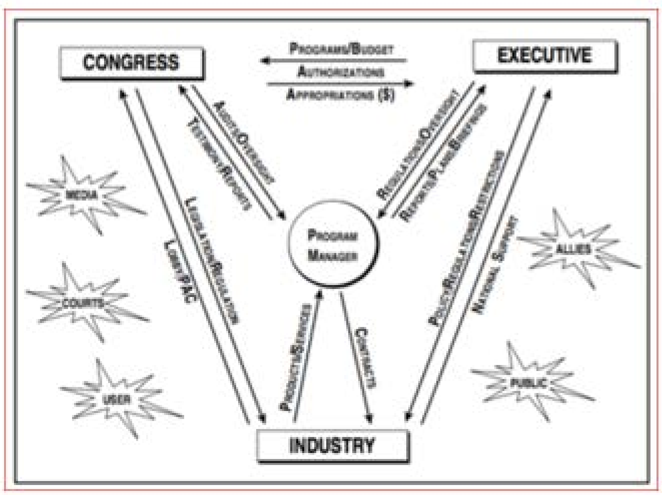


Figure 1. The Program Manager's Environment

By acknowledging the various types of interference involved in the program manager's execution of duties, we recognize why resistance to modifications of protocol can be so difficult to overcome. Once program managers discover how to navigate the difficult waters of program implementation, it seems understandable that they might resist modification of their proven methods of solving complex management issues. Additionally, by utilizing Augustine's Laws of military acquisition theory (Augustine, 1997), we can begin to understand the difficulty surrounding program managers and the frequently shifting environments in which they are expected to perform their duties. In addition to having a budgetary status that is uncertain and often-changing, many program managers will never see the completion of their own projects, as the life cycle of program managers is shorter than the average lifespan of a program, often even with the narrowest of perspectives.

Defense Acquisition Management Systems

Conducting a shortened analysis of the Defense Acquisition Management System is a difficult task, in part due to the sheer volume of steps, and in part due to the degree of acronyms utilized in conducting programs. DoD Instruction (DoDI) 5000.02 provides an outline of the in-depth protocols for conducting these processes, and it would also be useful to have *Introduction to Defense Acquisitions Management* (10th ed.) as a primer for understanding the colloquialisms and protocols involved in each step. However, it is important to acknowledge a few certain prime movers and processes in this process to begin any understanding:

- Defense Acquisition Executive (DAE)—The DAE is the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD[AT&L]). The DAE acts as the Milestone Decision Authority (MDA) for Major Defense Acquisition Programs (MDAPs) and Major Automated Information System (MAIS) programs.
- Milestone Decision Authority—The MDA establishes procedures for assigned programs using DoDI 5000.02, and tailors program strategies and oversight, including program information, acquisition phase content, timing, and scope of decision reviews and decision levels, based on the specifics of the product being acquired, including complexity, risk factors, and required timelines to satisfy validated capability requirements. The MDA is the sole and final decision authority.
- Program Acquisition Categories (ACATs)—All defense acquisition programs are designated by an ACAT (i.e., ACAT I through III) and type (e.g., MDAP, MAIS, or Major System)
- The Defense Acquisition Board (DAB)—The DAB advises the DAE on critical acquisition decisions when the DAE is the MDA. The DAE or designee will chair the DAB. An Acquisition Decision Memorandum (ADM) will document decisions resulting from reviews. Similar procedures will be established at the Component level for use by other MDAs.
- Program Managers—Under the supervision of Program Executive Officers (PEOs) and CAEs, program managers are expected to design acquisition programs, prepare programs for decisions, and execute approved program plans.

Figure 2 from DoDI 2010 gives an idea of what the acquisition phases and decision points might look like, so that readers might have an idea of processes involved in development of technology.

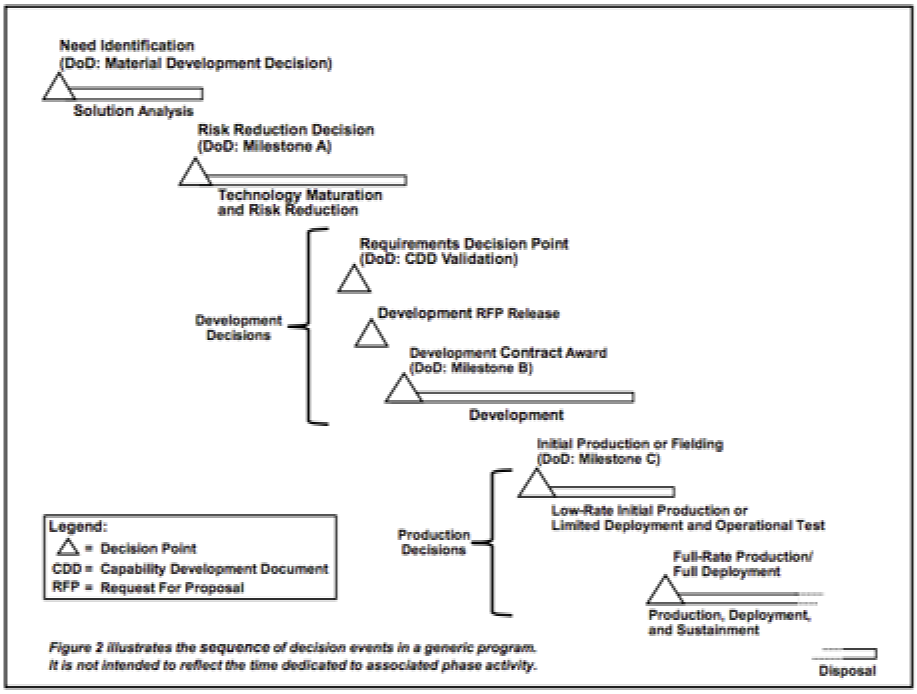


Figure 2. Generic Acquisition Phases and Decision Points

Opportunities for Advancement

It has long been proposed that, among other government practices, reporting data for federal acquisitions should be standardized to help make that data more accessible and useful. Recent research findings have shed significant light to support the importance of data quality. Accompanying such research progress includes advances in data analytics, data integration, data wrangling, and data visualization. Here *data wrangling* refers to any data transformation required to prepare a dataset for downstream analyses. Striding the entire process is the subject of data governance, which provides the centralized, and enterprise level oversight of corporate and enterprise data as an asset.

Many important acquisition management research issues have arisen from the emerging chief data officer practices. Additionally, acquisition planning, resource allocation, and other kinds of decisions depend critically on the data used in supporting these decision-making processes. Too often the question arises: How much do poor-quality data cost? How do untimely data, incomplete data, inconsistent data, untrusted data, and inaccurate data impact the eventual decision, and the subsequent operations and strategic making? Answering these questions will provide an acquisition guideline concerning how much it is worth investing to identify various root causes of poor-quality data, and continuously improve them throughout the acquisition decision cycle and their underlying data life cycle.

Research Approach

We propose a holistic enterprise approach to improving data capabilities for acquisition management, encompassing many interrelated research components:

1. A data platform with data technologies to handle a variety of data in high volume and velocity
2. Innovative data quality and data integration solutions, as well as state-of-the-art big data tools for improved data capabilities.

3. Data analytics ranging from simple business analytics to machine learning algorithms, to large scale math programing methods to improve acquisition management
4. Improved data capabilities in this data platform through emerging chief data officer and information quality research results and industry practices
5. Application of this holistic approach in various organizational settings to identify issues critical for future acquisitions research

Research issues will be addressed and research findings written for senior acquisition leaders and academic researchers. We expect three areas of research results:

- An assessment of the state-of-the-art, data-centric acquisition management practice
- Characterization of the salient features of successful outcomes
- A requirement analysis of tools, methods, and techniques that should be developed to improve acquisition management

We have used datasets from USASpending.gov to perform preliminary tasks. To begin with, we have downloaded 40 GB of DoD spending data, then loaded it into four tables that are categorized as PrimeAwardContracts, PrimeAwardsOFA, SubAwardContracts, and SubAwardGrants. The following are some preliminary findings:

- For some of the key fields, we have seen a number of data quality issues, such as misinterpretation, missing values, columns with no data, inconsistent representation, and fitness for use.
- There are chances of information dissemination when sensitive data is shared in public. It could potentially expose the information to users for exploitation purposes or hampering the business. This could be a possible weak link that is exposed here.

For instance, following is a sample use case that describes this scenario. For a company with prime_awardee_parent_duns : 217304393, we can easily retrieve the key information. The statistics of the company and the work that it does are exposed, which could be a possible risk. We can get information such as the following:

- We can see top products or services this company does for the DoD.
- We can also infer more information based on sub-awards by spending type or received ones.
- We can see more information on total funds awarded as prime and as sub-awardee.

(See https://www.usaspending.gov/transparency/Pages/RecipientProfile.aspx?DUNSNumber=217304393&FiscalYear=2017.)

In addition, when this information is combined with other information available on the Internet, more information might be inferred. To find information related to acquisition from the available data sources, we need to take a big dive into the datasets and see if we can design a model or logic to answer these big questions. We have begun to perform analysis to see how we can cross compare data from different units by applying Extract, Transform, and Load (ETL) and data analytics processes.

Specifically, we are replicating the same for other units to see how to operate in terms of resourcing, executing, deciding, and reporting the data well. Since this is big data problem, we plan to do a migration of all the data to Amazon Web Services to conduct further research analysis that traditional military acquisition theory and practice failed. Moreover, we are exploring opportunities on collecting more data for big data analysis for cross comparing the datasets from different units to see if we can infer relationships and conduct possible analysis to increase the business value.

Concluding Remarks

The mission of improving data capabilities for acquisition management is to provide leadership with timely access to accurate, authoritative, and reliable data supporting acquisition oversight, analysis, and decision-making. In this paper, we have reviewed the Defense Acquisition Visibility Environment, the evolution of acquisition policy, Augustine's Laws, related literature, and the defense acquisition management systems.

We explored opportunities for advancement in acquisition management and proposed a holistic enterprise approach to improving data capabilities for acquisition management, encompassing many interrelated research components. Next, we applied USASpending.gov datasets, unraveling data quality issues like misinterpretation, missing values, columns with no data, inconsistent representation, and fitness for use. We are poised to demonstrate that when sensitive data is shared in public, it could potentially expose the information to users for exploitation purposes or to hamper the U.S. acquisition management practice. Our research findings could strengthen U.S. acquisition decision-making processes while preventing adversaries from exploiting public data to hamper defense acquisition management practice.

References

Augustine, N. R. (1997). *Augustine's laws*. AIAA.

Brown, B. (2010). *Introduction to defense acquisitions managem*ent (10th ed.). Defense Acquisition University.

Cilli, M. et al. (2015). A systems engineering perspective on the revised defense acquisition system. *Systems Engineering, 18*(6), 584–603. doi:10.1002/sys.21329

Defense Acquisition University. *DAU Center for Defense Acquisition Research agenda 2016–2017*. Retrieved from http://dau.dodlive.mil/files/2016/01/ARJ-76_ONLINE-FULL.pdf

Gaither, C. C. (2014). Incorporating market based decision making processes in defense acquisitions. *International Journal of Defense Acquisition Management, 6,* 38–50.

Hagan, G. (1998). *Glossary: Defense acquisition acronyms and terms*. Fort Belvoir, VA: DoD, Defense Systems Management College.

Krzysko, M. (2012, February). The need for acquisition visibility. *Journal of Software Technology*, 4–9.

Krzysko, M. (2016). Acquisition decision making through information and data management. Retrieved from http://www.digitalgovernment.com/media/Downloads/asset_upload_file917_5737.pdf

Miller, A., & Ray, J. (2015, January). Moving from standard practices to best practices in defense acquisition. *Defense ARJ, 22*(1), 64–83.

Pennock, M. J. (2008). *Defense acquisition: A tragedy of the commons*. ProQuest.

Under Secretary of Defense for Acquisition, Technology, and Logistics (USD[AT&L]). (2007, November). *Operation of the Defense Acquisition System* (DoDI 5000.01). Washington, DC: Author.

Under Secretary of Defense for Acquisition, Technology, and Logistics (USD[AT&L]). (2015). *Operation of the Defense Acquisition System* (DoDI 5000.02). Washington, DC: Author.

Zhao, Y., MacKinnon, D. J., & Gallup, S. P. (2015, May). *Lexical Link Analysis (LLA) application: Improving web service to defense acquisition visibility environment.* Monterey, CA: Naval Postgraduate School, Acquisition Research Program.

Acknowledgments

Work reported herein has been contributed, in part, by Kiran Kumar Bandeli, Yupo Chan, Eva Lee, William Parsley, John Talburt, and many colleagues in the Department of Defense.

Appendix A: Defense Acquisition Visibility Environment (DAVE)

The Improving Data Capabilities for Acquisition Management Model, when fully realized, will feature three parts: the ***DAVE portal*, *DAVE Platform***, and AV Data Framework.

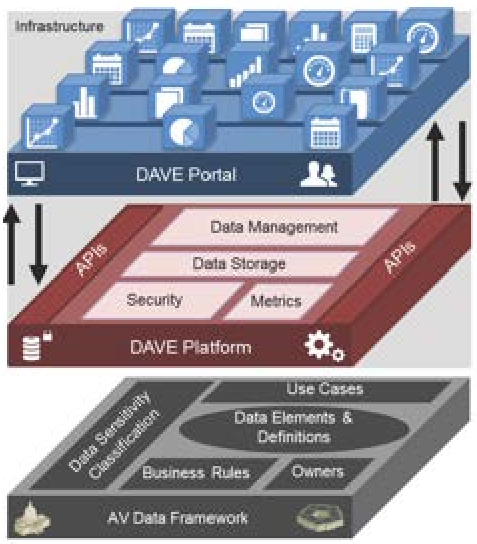


Figure 3. Improving Data Capabilities for Acquisition Management Model

DAVE Portal: The ***DAVE portal*** is a synthesis of interactive infrastructure including data visualizations, calendars, and project management tools that are set to continue to grow in scope and capability as DAVE expands. These diverse tools with analysis capabilities will help users answer such questions as, “Are we solving a business problem by assessing the efficiency and effectiveness of the project?” and “What value does the project add to acquisitions in the Department of Defense?”

DAVE Platform*:** The ***DAVE** platform* includes the Application Programming Interfaces (APIs) for data management, data storage, metrics, and security. The ***DAVE platform determines the APIs for facilitating data access, and determines to which party the information can be shared. This platform is made up of a single intuitive interface for all data, and supports the implementation of a data exposure strategy of acquisition of data, promotion data sharing, and also provides for flexibility. The APIs are the building blocks that allow for the integration of features or data, and the platform itself processes the data to get it to the state users require, as well as coordinating internal processes. These APIs also allow for a greater flexibility with analytics and faster development of new capabilities. The platform also facilitates data management, allowing a flexible yet technical approach for understanding, sharing, and protecting data. Throughout the entire process, APIs provide support for all the ***DAVE*** functions, ensuring security and proper access. At this time there are 350 structured item types and 84 unstructured document types, including data sets.

AV Data Framework: The AV Data Framework is the foundation on which the portal and platform are built and provides a number of essential elements including use cases, data elements and definitions, business rules, guidelines and markers regarding ownership of data, and data sensitivity classifications. The AV Data Matrix (AVDM) includes the definitions, definition owner, laws, regulations, data governance policies, data providers, and functions as the authoritative source for all data. It also includes data stewardship, representing the agreement and accountability for definitions and authoritative data. Currently, under ***DAVE*** governance, there are 582 acquisition data elements.

By providing department-wide access to DAVE to all branches of the armed forces, essential federal government agencies, Aerospace, Institute for Defense Analyses (IDA), the RAND Corporation, and the MITRE Corporation, among others, USD(AT&L) hopes to bring together all aspects of the acquisition process into one large data resource and to standardize that resource using entrenched data quality benchmarks and techniques. This data resource would then be made available to the appropriate recipients who could then take advantage of the value of timely and quality data. Governance would also be provided through the involvement of governmental agencies in the oversight and management of the data in DAVE. These agencies include the Assistant Secretary of the Navy (ASN), Information Systems (RDAIS), System Metric and Reporting Tool (SMART), Defense Acquisition Management Information Retrieval (System) (DAMIR), Defense Data Repository System (DORS), Defense Technical Information Center (DTIC/AIR), Earned Value Central Repository (EVCR), OUSD Budget Materials, Cost Assessment Data Enterprise (CADE), and other authorized entities. The input, oversight, and expertise of these groups would add their own aspects of value to the DAVE system and in return would be able to reap benefits of information on potential projects of their own.

The goals of this updated model of ***DAVE*** are clearly outlined in *Acquisition Decision Making Through Information and Data Management*: "streamline reporting, improve[d] availability of data for analytics, enabling decisions based on analytics through faster development of new capabilities, [and] incorporate[d] evolving security requirements" (Krzysko, 2016). The updated model of ***DAVE*** establishes a framework for improved and expanded support for the USD(AT&L), and does so by using data and the inherent value produced through increasing information quality to improve upon the current practices of programs managed by USD(AT&L). By continuing to develop and implement ***DAVE***, the USD(AT&L) will create a platform for big data, enable new acquisition capabilities, coordinate operational alignment, and support analysts to enable decision-making (Zhao, MacKinnon, & Gallup, 2015).

Analyzing Programs for Insight

Actions and projects undertaken in the U.S. Armed Forces are conducted in what is referred to as programs and are led by program managers. However, the exact definition of *program* varies among services. At the base level, we have programs, led by program managers, who are in turn led by Program Executive Officers, who are led by service level executives (Air Force, Navy), who are then led by the DoD-level leadership. What is missing from the following organizational chart are the five main goals of program managers, which tie together DoDI 5000.02 and ***DAVE*** goals: solve the business problem, solve the data problem, solve the organizational problem, drive efficiency, and drive effectiveness.

Figure 4. Program Organizational Chart

Appendix B: Evolution of Acquisition Policy

With the Better Buying Power initiative introduced in 2010 by then USD(AT&L) Ashton Carter, a new focus was directed to reworking how the DoD managed its complex acquisition practices. This model directly challenged the department to improve its methods of acquisition management, oversight, and process. This move was in part due to a budget that had risen to $1.7 trillion dollars, a 60% increase in under 10 years. USD(AT&L) was faced with a need to find a way to use the data at their disposal to change their practices as they came to expenditures and outcomes in order to develop more data-driven analytics and guidelines (Pennock, 2008).

> The message was clear: data and information were key to managing, overseeing and streamlining processes within the acquisition portfolio, but DoD would require diligence to obtain it. Data offered innovative perspectives on acquisition processes, delivering the necessary insight into acquisition cost, performance, affordability and other critical elements. Empowered with data, DoD leadership could report, analyze, and make informed decisions on the Department's complex acquisition portfolio.

This mandate then directed a team to focus on using structured data for insight into areas for improvement within the DoD, with areas of focus being data governance and using data as a service. The team could "identify authoritative sources of major acquisition information; have consistent, semantic definitions across the Department; measure data for accuracy, reliability and availability; and provide it to acquisition leadership for use in any visual tool giving them data-driven insight into the major acquisition portfolio." The pilot program proved that the DoD could manage and govern acquisition information and could provide data and information as a service in an efficient and effective way.

This was truly a game-changer for the USD(AT&L). Acquisition data was now understood as essential to effectiveness, and structured data was seen as the new way forward. This prompted the creation of "an on-demand environment that could provide data across the enterprise seamlessly and efficiently" (DAU, 2016). This became known as Acquisition Visibility (AV), which was made formally effective as of July 2009. *AV* is officially defined as "having timely access to accurate authoritative, and reliable information supporting acquisition oversight, accountability, and decision making throughout the Department for effective and efficient delivery of warfighter capabilities" (USD[AT&L], 2007). AV quickly grew into an essential source of information across the Department, with teams providing functional, technical, and data expertise. AV now includes all major defense acquisition programs and shares operating costs as well as earned value management data.

AV increases transparency, which aids reporting, helps reduce costs, and is responsive to users (USD[AT&L], 2015).

In July 2015, the DoD revised its defense acquisition system policy, moving from DoDI 5000.01, *Operation of the Defense Acquisition System*, to DoDI 5000.02. DoDI 5000.01 provided a basic set of definitions and three overarching policies that governed the defense acquisition system: flexibility, responsiveness, and innovation. Part of the cause for this shift in policy was a need to address suspected root causes hindering higher success rates. DoDI 5000.02 established a management framework for translating mission needs and technological opportunities into "stable, affordable, and well-managed acquisition programs" (Brown, 2010). DoDI 5000.02 established a general approach for managing all defense acquisition programs while authorizing program managers and the Milestone Decision Authorities (MDAs) discretion to exercise prudent business judgment in structuring tailored, responsive, and innovative programs (Brown, 2010). DoDI 5000.02 placed increased emphasis on the use of systems engineering activities applied early in the project life cycle, so that meaningful tradeoffs between capability requirements and life-cycle costs could be explored and to ensure that realistic program baselines were established such that associated life-cycle costs would fit within future budgets (Cilli et al., 2015).

> This effort to move away from the open-loop capability requirements writing approach toward a closed-loop capability requirements writing process informed by rigorous assessments of a broad range of system level alternatives across a thorough set of stakeholder value criteria to include life-cycle costs, schedule, and performance. (Cilli et al., 2015)

As of 2016, the OSD is fed reports from a variety of unaffiliated data sources. Individual program offices are responsible for managing and streamlining their own programs, and are given the authority to make modifications to their programs in ways that they best see fit. Unfortunately, this practice creates a multitude of largely unstructured, loosely-governed data that are difficult to manage, report on, or standardize, which are then fed to the OSD without first having data quality best practices applied (Gaither, 2014). Acquisition decisions are primarily made at the service level, which is understandable considering they are the parties who will be responsible for said items in the field, but this division between program managers, MDAs, and the OSD has created communication gaps which need to be overcome in order to adequately manage and translate data from one entrenched group to another. In the past, the USD(AT&L) has expressed that the following are areas where improvement is essential to the continued success of Improving Data Capabilities for Acquisition Management, and acquisitions in general (Hagan, 1998):

- Initial operational test ratings
- Incorrect testing and management of program expectations and deliverables
- MDAP Research Development Test & Evaluation (RDT&E) funding growth from original baselines
- Falling competition rates
- Subcontracting roadblocks
- Overly optimistic program baselines
- Lower development schedule growth compared to development cost growth

Users also need to retire legacy reporting systems, but still must report their data consistently before, during, and after system retirement takes place. This conflicts in principal with the DoD goal of encouraging deeper data analytics by confusing data types and targets (Miller, 2016). Currently, acquisition data management functions by combining data and information access, federated data stores, and a variety of older data resources (DAMIR, KScope, AIR, Data Matrix, etc.) into one large data repository. Improving data capabilities for acquisition management seeks to provide the DoD with data and analysis support capabilities to better inform the acquisition community. At its fully-realized potential, improving data capabilities for acquisition management would function as the location, platform, and framework for the DoD to access and utilize this data more fully than ever before. This newly revised model of Improving Data Capabilities for Acquisition Management would represent a shift from the collection of capabilities into one fully integrated and mature analytics system, where an integrated data processing background would support an agile environment and efficient data and information access.

Appendix C: Augustine's Laws and Major System Development Programs

In 1979, Norman Augustine, the then assistant director of Defense Research and Engineering in the Office of the Secretary of Defense, penned a tongue-in-cheek piece on the pitfalls of military acquisition theory and practice, and to this day, it is touted as a highly accurate, if comical, display of the discouraging practices involved with acquisitions in the U.S. military. All jokes aside, and there are many jokes, his analysis was accurate in that even in 1979, Augustine could accurately predict the degree to which military spending, employment of civilian population, time management decline, and program failures would continue to the modern military procedures post 9-11. His insights were so striking because, first, he was in the position to make them, and second, unlike so many of his predecessors and even successors, he was frank about these problems and their sources. To attempt to construct a better model for the acquisition programs would be impossible without incorporating many of Augustine's "Laws." Some of the more appropriate ones reduced to theories for incorporation include the following:

- The bottom half of the production produces less than 20% of the output.
- Delivery of items will take on average one-third more time than initially estimated.
- The "doing" time has not increased, but instead the "planning time" has.
- Systems are now obsolete almost before they enter the field.
- In non-competitive processes, time expands to fit the work prescribed.
- "Lightning" in the form of unforeseen circumstances, usually negative (or unknown unknowns, as compared to known unknowns) will strike every project, but the cost-cutting bidding measures that prevail in cost-reimbursing contract work doesn't allow for controlling or budgeting for said factors.
- More complex systems are always more expensive, but don't always translate into contributing that much more success of military actions in the field, and especially not to the degree to which they are more costly.
- Most programs get a one-year honeymoon period, and from there the chances of being cancelled increase every year by a linear factor.
- Price-reduction bidding incurs the problem of rewarding a contract to a new business who does not understand the difficult lessons learned by the original producer of the item, who set the original price that began the bidding process.
- Congress will approve the defense budget for the given year as: the budget of the prior year, plus 3/4ths of what is requested, and minus a 4% tax.
- Regulations as a management surrogate will grow at an exponential pace.

- Program managers responsible for long-term projects often are not in their position long enough to see most of their project completed.
- "By the time the people at the top are ready for the answer, the people at the bottom have forgotten the question."
- There is no incentive system to assist in rewarding good managers, and vice versa.
- It would be pertinent to reduce the number of acronyms used to clear up understanding.
- Software is always expanding and increasing in complexity.
- If you send money to the management of a project that is in trouble, they will remember you the next time they need money.
(Augustine, 1979, 2015)

Appendix D: A Preliminary Literature Review

The following are theories that offer critiques of the current system and opportunities for constructive modifications.

Lexical Link Analysis

In 2015, researchers from the Naval Postgraduate School utilized Lexical Link Analysis (LLA) as a way of improving web services for Improving Data Capabilities for Acquisition Management and found there were significant opportunities for further research (Zhao et al., 2015). LLA, a hash-like process, was used to find a "fit" between budgets, final products, and requirements using reports, visualization, and linguistic analysis. Collaborative learning agents for pattern recognition were also tested, and may allow for scaling up to big data. Topics mentioned as opportunities for further study were system self-awareness, big data architecture and analytics, and deep learning. By examining acquisition data sources, we might be able to perform big data analytics and gain business insights from contractor relationships, budget analysis, time series analysis, and so forth. Additionally, system self-awareness might be utilized to compare behavior among nodes, and compare these relationships to business processes.

"Push, Practicality, and Pull" Theory of Standardization of Practices

In *Moving From Standard Practices to Best Practices in Defense Acquisition*, Alex Miller and Joshua L. Ray (2015), economics professors from University of Tennessee, Knoxville, and members of the Defense Acquisition University (DAU), looked at utilizing "Big Checks" as a method to help illustrate cost-saving and value-production to communicate value in data-driven investments. They found that personal-best-interest was the primary motivating factor in defense contractor work, and suggest extending this model to defense logistics and acquisition. Their prime theme, "What's in It for Me?" (WIIFM), found that the following six forces work collectively to influence the extent to which organizations are able to turn isolated best practices into widespread standard practices:

- Inherent Stakes
- Making Advantages Visible
- Replicability of Work
- Implementing Standard Work
- Organizational Alignment
- Driving Compliance

Producing a single communication system could help break through the negative cycles of acquisition times, budgetary oversights, and data governance issues, if only the ideas were shared appropriately. This stands as a model for the prime mover of Improving Data Capabilities for Acquisition Management as well as with WIIFM. In both instances, the stakes need to be visible and clear. For example, efforts to reduce acquisition cycle times produced impressive breakthroughs, often with cycle times reduced 40–60%. And yet, there is little evidence that the efforts producing these performance gains are encouraged as standard practices:

> Consider the perspective of members of a defense acquisition program team who had greatly reduced their source-selection time, allowing a badly needed system to be put under contract months earlier than expected. No one on the team could identify a single request to share ideas with other source-selection teams. Furthermore, members of the successful team were not confident that members of this team would apply lessons learned from their effort, even to their own future source selection work! (Miller & Ray, 2015)

Defense acquisition suffers from what can only be described as an abundance of, and yet a severe drought of, communication in policy execution. Often, the process that is undertaken to accomplish a program takes so long that by the time the higher brass have decided about an appropriate solution, there is no longer a problem at the lower ranks. And yet, it is in these same situations that DAVE could function as a model to increase communication in the acquisition process through data management strategies. It seems that the current model of defense communication does not make a large enough effort to systematically share best practices, even though the work performed across departments is very similar and often utilizes the very same contractors.

> The most common answer in response to questions about this lack of standardization was very revealing in that it highlighted the importance of perceived high stakes as a driver: "Standardization across organizational boundaries is hard. Why do it if we can get satisfactory performance working on our own?" … The perceived stakes inherent in defense acquisition are not sufficiently high to be an important driver of efforts to standardize and replicate processes. Note the emphasis on perceived stakes; the actual stakes are really quite high, suggesting the need for managers to make the stakes more visible. (Miller & Ray, 2015)

Tragedy of the Commons

In his research paper, *Defense Acquisition: A Tragedy of the Commons*, Michael Pennock (2008) argued that the DoD should recognize the Tragedy of the Commons as it relates to the development and implementation of new military contractor work and pursue mature technologies as project and programs expand in scope. Pennock presents the reader with a model with mathematical analysis for the conundrum and consequences of increased project requirements burdening the system by adding increasingly expanding scope and subsequently immature technologies to meet uncharted project scope areas. This process unnecessarily burdens the system and eventually causes failure, budgeting crisis, and cost-overages.

> To understand this situation, a mathematical model of a series of acquisition programs is developed and analyzed. It reveals that when differing stakeholder interests come into play, the program suffers from a classic tragedy of the commons. The program serves as a common resource for these stakeholders, and they are incentivized to pursue aggressive

> performance requirements that necessitate immature technology. The critical aspect of this result is that this behavior is rational. In other words, the behavior we see is exactly what we should expect to see. This suggests that the recent trend in defense acquisition to reduce costs by aggregating the requirements of multiple groups of users into a single program may actually be counterproductive. This result has implications for the policy makers, managers, and engineers that are responsible for developing and deploying defense systems. (Pennock, 2008)

Big Data

When we think about big data, we typically think of the "Big 3 Vs": velocity, volume, and variety, and with the "4 Big Questions" (Hagen, 2015)

- Where will big data and analytics create advantage?
- How should we organize to capture the benefits of big data and analytics?
- What technology investments can enable the analytics capabilities?
- How do we get started on the big data journey?

When considering the value of these questions in terms of the acquisition process defined thus far, it is essential to consider the purpose and value of supply chain data structures. If Big Data can get a handle on the vast amount of resources at the behest of the USD(AT&L), the ability to move and sort not only data, but actual goods and services to areas of need could be monumental. If every good, service, and data point was given a stock keeping unit (SKUs), and these SKUs were consistently measured and accountable using our Big Data resources, in the event of a need in one area, the SKUs could automatically be routed to that area as a sort of economic triage immune system response. Instead of being bombarded with white blood cells or antibodies, however, a program manager in need of lumber might get a notification that three program managers have excess or unused lumber at the moment, and these could be selected by geographic distance to find the best logistical match. And unlike typical supply chain structures, the U.S. military is in a unique position to control the goods, services, supply chain, and communication devices relative to its operations. By incorporating already in owned transportation techniques to ease transport and arrival, USD(AT&L) could reap the rewards of a supply chain windfall, not unlike how Wal-Mart based much of its low-price strategy on its ability to ship the predetermined number of goods to designated stores by using single palettes for multiple good types. This could also help to build a predictive model so that the next time the program manager is almost out of lumber, the model will already be working on the best possible solution before the need becomes a reality.

Map/Reduce & Scan/Hash

Two final non-mutually exclusive suggestions stem from reading thus far. First, a Map/Reduce algorithm could be applied to create associations with terms, groups, and contractors in an attempt to learn from previous work orders, experiences, reports, and so forth. This would allow us to create a working and searchable knowledgebase that is responsive in real-time to inputs in Improving Data Capabilities for Acquisition Management, and generates a report of potential helpful pieces of past reporting (which is sensitive to security needs, of course). If a program manager were entering in his cost reporting data for his program for XYZ Manufacturing Company, Improving Data Capabilities for Acquisition Management could point the program manager to alternative contracts with XYZ Mfg. This could help not only expedite work order forms and billing information, but might also allow the program manager to get into touch with individuals in other departments who have experience working with XYZ Mfg. This manager could utilize ***DAVE*** to then quickly message the other party to ask questions, to ask for tips, or to request insight into their experience with XYZ.

A Scan/Hash function that searched for acronyms and replaced them with complete and readable terms for use in Improving Data Capabilities for Acquisition Management might save many users headaches and help to encourage simplification, if only at the linguistic level. This could create ease of use solutions for users, and a general increase in understanding of complex acquisition processes.

Determining New System Design Requirements to Optimize Fleet Level Metrics Under Uncertainty

Satadru Roy—is a PhD student of Dr. William Crossley in the School of Aeronautics and Astronautics at Purdue University specializing in the field of Aerospace Systems. His research interests center on design optimization and systems engineering. [roy10@purdue.edu]

Navindran Davendralingam—is a Research Scientist in the School of Aeronautics and Astronautics at Purdue University. He received his PhD from Purdue University in Aerospace Engineering in 2011. He is currently conducting research in the Center for Integrated Systems in Aerospace (CISA) led by Dr. Daniel DeLaurentis. [davendra@purdue.edu]

William A. Crossley—is a Professor in the School of Aeronautics and Astronautics, Purdue University, where he has been on the faculty since 1995. His research and teaching interests focus on design and optimization of aerospace vehicles and systems of systems. [crossley@purdue.edu]

Parithi Govindaraju—is a PhD student of Dr. William Crossley in the School of Aeronautics and Astronautics at Purdue University specializing in the field of Aerospace Systems. His research interests center on design optimization and systems engineering. [pgovinda@purdue.edu]

Abstract

Traditional approaches to design and optimize a new system often do not consider how the operator will use this new system alongside the other existing systems. This “handoff” between the designs of the new system and how this new system operates with the group of systems leads to the sub-optimal performance of the new system when measured with respect to system-level objective. Aircraft design choices made to meet a set of requirements dictate the performance of the aircraft, and the aircraft performance influences how the operator might use the aircraft. Further, the presence of uncertainties in predictions of the new aircraft performance and costs and uncertainties in the amount of payload to transport further exacerbate the problem of determining these requirements. Recent efforts have posed approaches to address this problem, but generally with a deterministic perspective. This research adopts a previously developed subspace decomposition approach and integrates features from robust/reliability based optimization to address the uncertainties and solves two application problems—a military and a commercial airline application. The result demonstrates the ability of the framework to identify the design requirements for the new aircraft, and a posterior analysis indicates that the framework acceptably handles the uncertainties.

Research Issue

The *Better Buying Power 3.0* document (Kendall, 2014) states, “Defining requirements well is a challenging but essential prerequisite in achieving desired service acquisition outcomes.” Traditional acquisition processes focus on development of requirements at the system-level. Current acquisition analyses of design alternatives are disjointed from considering operations (the way an end user operates these new systems alongside existing ones), resulting in inefficiencies at the higher aggregate level (Taylor & Weck, 2007; Mane, Crossley, & Nusawardhana, 2007). Typical design practice for new systems assumes a “handoff” between the design of the new, yet-to-be introduced system, and the operations on how the system impacts top-level performance.

The authors proposed an approach that would include top-level requirements for a new system as decision variables in an optimization problem. With the objective to maximize (or minimize) a fleet-level performance metric, then an optimization algorithm should determine the “right requirements” as part of finding the optimal set of decision variable values. Using aviation examples, one can pose the optimization problem that included top-level requirements as decision variables along with new system design variables and operational decision variables. The resulting formulation is a mixed-integer nonlinear programming problem that is very difficult if not impossible to solve in reasonable time. The authors and their colleagues have developed a decomposition approach that allows solution of this problem, with a few minor modifications from the original problem.

The initial efforts concentrated on demonstrating that solving the decomposition approach was practical and that the results were useful; however, those initial efforts could not address data uncertainties in the problem. The recent work has identified and demonstrated how to include consideration for various types of data-driven uncertainties as well. With the focus on aviation examples, the work first considered an application of the decomposition approach under uncertainty to military air cargo transportation using actual data from the U.S. Air Force Air Mobility Command (AMC) as the basis for a set of example problems. Then, to explore the flexibility of the decomposition approach under uncertainty, data from the Bureau of Transportation Statistics provided the basis for another set of example problems representative of commercial airlines.

This paper presents how the approach applies to both military air cargo problems and to commercial airline problems and how the approach handles uncertainties in the aircraft design sub-problem, propagates those uncertainties to the allocation (commercial airline) or assignment (military air cargo) sub-problem, and additionally considers demand uncertainty in the allocation or assignment sub-problem. While the overall decomposition framework can address these two different aviation problems under uncertainty, there are some specific modifications necessary to represent these two different problems.

The approach is able to identify the best requirements for a new aircraft for both the commercial airline and military air cargo problems. *A posteriori* analysis of the resulting design shows the advantages that the approach under uncertainty has over deterministic approaches to the same problems.

Subspace Decomposition Approach

This section describes in details the methodology that uses the previously developed subspace-decomposition approach (Mane et al., 2007; Govindaraju, Davendralingam, & Crossley, 2015). The approach serves as a ‘meta-algorithm’ framework within which specific choices in performance metrics and resource constraints can be made for each of the two problem instantiations we have solved (AMC and Commercial Airline) in prior work (Roy et al., 2017; Govindaraju et al., 2015). The description of each subspace and the information flow between subspaces appears in Figure 1.

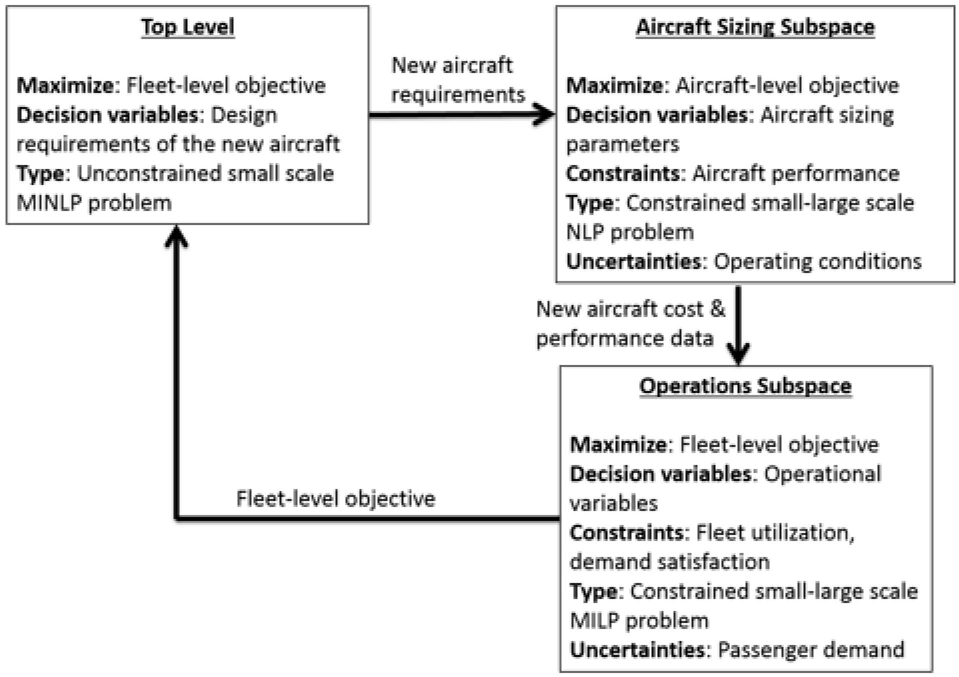


Figure 1. Overview of the Sequential Decomposition Framework

Top-Level Subspace

The top-level problem seeks to maximize the fleet-level objective of the operator, based upon the choice of the design requirement of the new yet-to-be-designed aircraft. These top-level requirements include design range, payload-carrying capacity, etc. of the new yet-to-be-designed aircraft. This level is a small-scale Mixed Integer Non-Linear Programming (MINLP) problem and is solved either using an MINLP solver or by performing a pseudo enumeration.

Maximize Fleet level objective

Design variables Design range   **Top: MINLP**

Payload carrying capacity

Aircraft Sizing Subspace

The decision variables from the top level appear in the aircraft sizing sub-space as parameters. Starting from these top-level requirements, this subspace solves an aircraft design optimization problem with the objective that minimizes the design mission direct operating cost. The decision variables for this sub-problem are the variables that defines the wing geometry such as aspect ratio, taper ratio, sweep, etc. and the engine parameters like static thrust, bypass ratio, fan pressure ratio, and so forth. Further, the portion of the aircraft conceptual design phase known as “aircraft sizing,” usually uses empirical equation and simplified physical models to predict the cost and performance of the aircraft. The limited knowledge available at this phase of the design process combined with the modeling fidelity results in high uncertainty.

Minimize Design mission expected direct operating cost		
Design variables Wing design variables		
Engine design variables		**Size: NLP**
Subject to Aircraft performance constraints (using RBDO)		

For instance, an aircraft is sized for its design mission based on a set of nominal values for operating conditions (e.g., cruise altitude). However, when evaluating the operating missions to determine block time and fuel consumed on the flight, there might be a variation in winds aloft, which would alter the block time and fuel consumed. Additionally, predictions of the aircraft performance and characteristics, like parasite drag, that use low-fidelity models will have associated uncertainty. It is therefore necessary to simulate the effect of uncertainties on the design parameters, in the absence of closed form mathematical expressions, for subsequent inclusion in the resulting aircraft sizing optimization problem. We employ a reliability-based design optimization (RBDO) formulation on the new aircraft that is subject to a collection of uncertain parameters. This sub-problem is a Non-Linear Programming (NLP) problem that can be solved using a choice of NLP solver such as the *fmincon* function in MATLAB.

Maximize Expected Fleet-level objective		
Design variables Allocation (integer type)		
Payload (continuous type)		**Alloc: MILP**
Subject to Aircraft utilization constraints (nominal and worse case)		
Demand constraints		

Operations Subspace

Operations subspace seeks to solve how the operator uses the new yet-to-be-designed aircraft alongside the existing fleet of aircraft. This is an allocation problem that allocates the new aircraft together with the existing aircraft with the goal to maximize the fleet-level objective. The strategy involves assigning or allocating the fleet on various routes. This sub-problem is posed as a Mixed Integer Linear Programming (MILP) problem with both integer (allocation variables) and the continuous (payload) type variables and is solved using the CPLEX solver available within the GAMS (Brooke et al., 1998) software package. This sub-problem is subjected to operational constraints such as aircraft utilization, demand, and so forth. Further the demand in this subspace is uncertain. The amount of payload to carry across the various routes is an uncertain parameter. Thus, we have two levels of uncertainties that interact and need some strategies to address the propagation of uncertainty from one domain to the other. The new aircraft coming out of the aircraft sizing subspace has uncertain performance and cost coefficients. Our approach employs an Interval Robust Counterpart (IRC; Lin, 2014) formulation to address this uncertainty propagation from the sizing sub-space to the allocation subspace. We size the aircraft at two cases of the uncertain parameters of the aircraft sizing subspace: a nominal case and a worse case and use the IRC formulation to enforce the worse-case performance and cost in the allocation constraints using some tolerance limit. An overview of the operations sub-problem (Alloc: MILP) appears below.

In the following two sections, we detail application of the subspace decomposition approach for the case of setting optimal requirements for military air cargo, and, for commercial airline systems. We mainly highlight key differences in the modeling approach for each subsection to illustrate flexibility of the framework in accommodating unique problem characteristics of each case.

Applications of Subspace Decomposition Approach

Case 1: Military Air Cargo

The subspace decomposition approach in the prior section is used to determine the optimal requirements of a new, yet-to-be introduced system (here, strategic airlift aircraft), which will operate alongside other strategic military airlift aircraft of the United States Air Force Air Mobility Command (AMC). The problem was motivated by the USAF AMC's emphasis on reducing fleet-wide fuel consumption. The objectives are to maximize expected fleet productivity and minimize expected fuel consumption. As these are competing objectives, the problem is posed in a multi-objective sense where fleet-wide fuel consumption is minimized and a minimum acceptable fleet productivity level is set as a constraint that is varied to generate a series of non-dominated Pareto solutions. Data on cargo demand is obtained from the Global Air Transportation Execution System (GATES) dataset for the year 2006. Figure 2 illustrates the subspace decomposition of the AMC problem statement.

Differences in Top Level Subspace

The top-level optimization problem does not include any nonlinear constraints and only has bounds imposed on the top-level decision variables. Equations (1) to (4) describe the deterministic formulation of the top-level problem; the formulation incorporating uncertainty appears in latter subspaces.

$$Minimize \quad Fleet\ fuel\ (Pallet_x, Range_x, Speed_x) \tag{1}$$

$$Subject\ to \quad 14 \leq Pallet_x \leq 38 \quad \text{(Design pallet capacity bounds)} \tag{2}$$

$$2400 \leq Range_x \leq 3800 \quad \text{(Range at max. payload bounds in nmi)} \tag{3}$$

$$350 \leq Speed_x \leq \quad \text{(Cruise speed bounds in knots)} \tag{4}$$

$$Pallet_x \in Z^+ \quad Range_x, Speed_x \in R^+$$

Equation 1 describes the objective function that seeks to minimize the fleet-level fuel consumption using pallet capacity, range and cruise speed of the new, yet-to-be-introduced aircraft type X as decision variables. Equations 2–4 describe the bounds for the top-level design variables. The values for the bounds were based on strategic airlift requirements, and characteristics exhibited by current cargo transport aircraft (Gertler, 2010; Graham et al., 2003). Here, the design requirement decision variable describing payload capacity uses an integer number of pallets, while the design range and design speed decision variables are continuous.

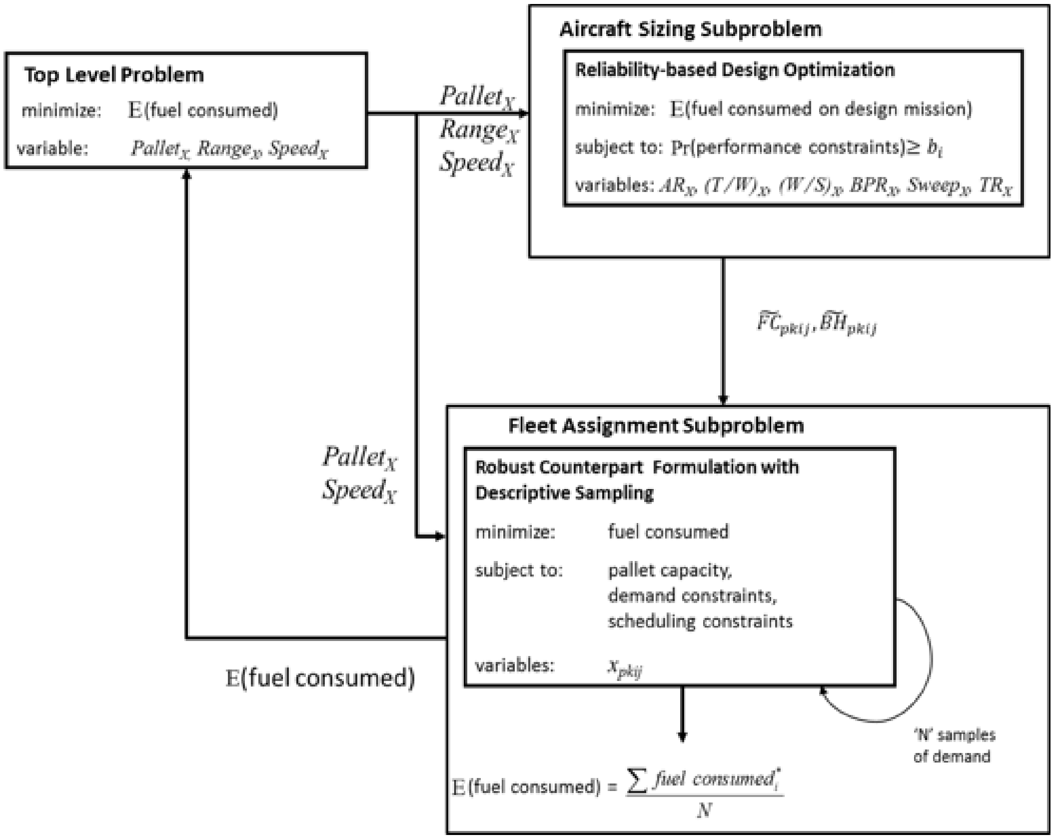


Figure 2. Subspace Decomposition strategy for the USAF AMC Application

Differences in Aircraft Sizing

Uncertainty in Design Parameters

The conceptual phase of the aircraft design process relies upon semi-empirical equations and simplified physics models. The limited knowledge available about the system definition at this phase of the design process combined with the usage of low-fidelity modeling tools results in high uncertainty. Aircraft sizing typically determines the size, weight and performance of an aircraft to meet its design mission based on a set of nominal values on operating conditions (e.g., cruise altitude). However, when evaluating the operating missions to determine block time and fuel consumed on the flight, there might be a variation in assigned altitude, routing, speed, and so forth, which would alter the block time and fuel consumed. For instance, there is uncertainty in the prediction of the parasite drag coefficient. In this example, a scaling factor k_{C_D} follows a distribution to represent the uncertainty in the parasite drag prediction, so that the "actual" coefficient relates to the "predicted" coefficient in the following manner:

$$C_{D_0\,actual} = k_{C_D} \text{x} (C_{D_0\,predicted})$$

To address the uncertainty related to operations and predictions of the new aircraft performance in the aircraft sizing subspace with reasonable computational expense, the Analysis of Variance (ANOVA) technique, a sensitivity analysis method, determined the subset of the most important parameters that influence the outputs under consideration

(Montgomery, 2008). This investigation assumes triangular distributions for the scaling factors of identified parameters listed in Table 1.

Table 1. Triangular Distributions of the ANOVA Identified Uncertain Parameters in the Aircraft Sizing Subspace

Uncertain Parameters (ξ)	Lower limit	Mode	Upper Limit
C_{D_0} multiplier, k_{C_D}	0.90	1.0	1.10
Specific Fuel Consumption, SFC [hr^{-1}]	0.45	0.5	0.55
Oswald efficiency multiplier, k_{e_0}	0.95	1.0	1.05
Cruise altitude [ft]	32000	35000	38000
Pallet mass [lbs]	7200	7500	7800

The aircraft sizing sub-problem seeks to minimize the fuel consumption of the new, yet-to-be-introduced aircraft for the values of design range ($Range_x$), pallet capacity ($Pallet_x$), and cruise speed ($Speed_x$) from the top-level problem. With the top-level objective to minimize fleet-level fuel consumption and the aircraft-sizing objective to minimize the fuel consumed by the new aircraft for its prescribed design range, pallet capacity, and cruise speed, a slight disconnect exists between the objectives of these two levels. The difference in the objectives is that at each aircraft sizing iteration the minimization of fuel consumption uses a single combination of fixed values for design range, pallet capacity, and cruise speed—this is the typical case in aircraft design where these quantities are set as requirements for some “representative design mission.” However, the top-level optimization problem drives the question of “What requirements do we need to set in the first place?” by searching through the decision space of the top-level variables to find aircraft requirements that optimizes fleet-level operational aspects of how the aircraft is used.

For example, consider the dimension of design range—as the top-level problem searches across values of range, this naturally changes the set of feasible routes that the new aircraft can fly, thereby changing how the fleet comprised of existing and new aircraft serves the overall route network. By doing so, the top-level problem seeks additional fleet-wide fuel savings that these operational aspects reflect as a function of the decision variables. Therefore, the aircraft sizing objective can be viewed as a subset of the top-level problem objective. Because the type of aircraft assigned on individual flight segments drives the total amount of fuel consumed by the fleet, an aircraft designed for minimal fuel consumption will lead to improved fleet utilization that reduces fleet-level fuel consumption when compared to fleet operations using only the fleet of existing aircraft. The approach in this work poses the aircraft design sub-problem in the context of Reliability Based Design Optimization problem to account for uncertainty in the design phase.

The Reliability-Based Design Optimization (RBDO) formulation (shown below) represents the aircraft design under uncertainty problem.

$$\underset{x}{\textit{Minimize}}\ \mathrm{E}\left[f(x,\xi)\right]$$
$$\textit{Subject to}\ \mathrm{P}\left[g(x,\xi)\leq 0\right]\geq b_i \qquad \forall\ i = 1, 2, \ldots, m$$

x : set of decision variables

ξ : set of uncertain parameters

b_i : desired probability of satisfying the i^{th} constraint

Aggregating the outputs for each realization (sample) of the uncertain parameter allows for the estimation of statistical measures such as expectation and probability, which the objective and constraint function evaluations require. The objective of the aircraft sizing subspace is to minimize the fuel consumption of the new aircraft X using the decision variables listed in Table 2. For each function evaluation of the top-level problem, the current values of $Pallet_x$, $Range_x$, and $Speed_x$ become fixed parameters for the aircraft sizing problem. Table 2 summarizes the decision variables, uncertain parameters, and constraints in the aircraft sizing optimization problem.

Table 2. Decision Variables and Constraint Limits in the Aircraft Sizing Optimization Problem

Decision Variables, (x)	Lower Bound	Upper Bound
Wing Aspect Ratio, AR_x	6.00	9.50
Thrust-to-weight Ratio, $(T/W)_x$	0.18	0.35
Wing Loading [lb/ft^2], $(W/S)_x$	65.00	161.00
Engine Bypass Ratio, BPR_x	4.50	14.50
Wing Leading Edge Sweep [deg], $Sweep_x$	10.00	35.00
Wing Taper Ratio, TR_x	0.10	0.40
Constraints	**Value**	
Takeoff Distance [ft]	≤ 8500	
Landing Distance [ft]	≤ 5500	
Second segment climb gradient	≥ 0.025	
Top-of-climb rate [ft/min]	≥ 500	

The aircraft sizing sub-problem includes performance constraints such as limits on takeoff and landing distances and upper and lower bounds for the decision variables. The RBDO formulation optimizes the expected performance metric of interest and ensures that the probability of satisfying the performance constraints is greater than or equal to the user-defined reliability level, b_i, considering the uncertainty present in this sub-problem.

Differences in Fleet Operations

This subspace mathematically represents the AMC's operations where the AMC fleet flies cargo missions to deliver pallets of supplies on an "as-needed" basis without a predetermined, long term schedule. The fleet allocation model here considers the multiple destination nature of the flight path for each aircraft, where an aircraft may fly from point A to B and then on to C—this in contrast is different to the airline case where airline aircraft are assigned to fly back and forth on specific segments points. This multiple destination travel path prompts the need to include tracking of tail numbers in the fleet operations subspace. Furthermore, the unscheduled and uncertain nature of demand for cargo transportation includes unknown origin and destination pairs of trips as well—this is modelled using random sampling of starting points for aircraft where the random sample mimics the end of the previous day's flight termination point of the aircraft. The Interval Robust Counterpart (IRC) formulation addresses uncertainty in parameters within AMC fleet operations model; in this case the uncertainty associated with the fuel consumption rate $\widetilde{FC}_{p,k,i,j}$, and in the flight block hours $\widetilde{BH}_{p,k,i,j}$, on given routes in the service network. The optimization problem of the fleet operations model seeks to minimize the fleet-level fuel consumption while enforcing a constraint on productivity.

Case 2: Commercial Airline

We apply the subspace decomposition approach, as a modified version of the AMC case, to the case of a commercial airline application. These modifications arise from the statistical differences in cargo demand between the AMC case study and passenger demand for a commercial airline and from the underlying business model where airlines will set and publish a schedule from which the traveling passengers select flights and purchase tickets. The highly uncertain nature of demand in the AMC case, versus the more symmetric and seasonal nature of demand in commercial applications, prompts different computational strategies within the approach presented here. The detailed subspace decomposition framework for the commercial airline application appears below (also appears in Roy et al., 2017). For the commercial airline application, the airline operation subspace is further sub-divided into two subspaces—airline allocation and a profit evaluation block.

Top-Level Subspace

The top-level optimization problem for the commercial airline application, seeks to maximize the expected fleet-level profit of a representative airline based on the choices made about the design requirements for the new, yet-to-be designed aircraft; here, the range and passenger seating capacity are the design variables in this top-level problem. Like the AMC formulation, the top-level optimization problem is unconstrained except for bounds imposed on the decision variables. The following equations describe the formulation of the top-level problem; consideration for uncertainty, as reflected in the expectation of profit appears later in the aircraft sizing and airline operations subspace.

$$
\begin{aligned}
\text{Maximize}_x: &\quad E[Fleet\ Profit] \\
Subject\ to: &\quad 75 \le SeatCapacity_x \le 250 \\
&\quad 500 \le Range_x \le 2600 \\
&\quad SeatCapacity_x \in \mathbb{Z}^+, Range_x \in \mathbb{Z}^+
\end{aligned}
$$

The objective function here seeks to maximize fleet-level profit using passenger seating capacity and range of the yet-to-be-introduced aircraft type X as decision variables. The two constraints describe the bounds for the top-level design variables of aircraft passenger seating capacity and range. The values for the bounds on these design variables were based on typical characteristics of current class of aircraft. Here, the design requirement decision variable describing passenger seating capacity and the design range are both integer variables. While the expectation term appears in the objective function of the top-level formulation, the source of uncertainty associated with the expectation term comes from the aircraft design and fleet allocation subspaces. Our discussion in these latter sections will make clear the evaluation of the expectation term for the top-level objective function.

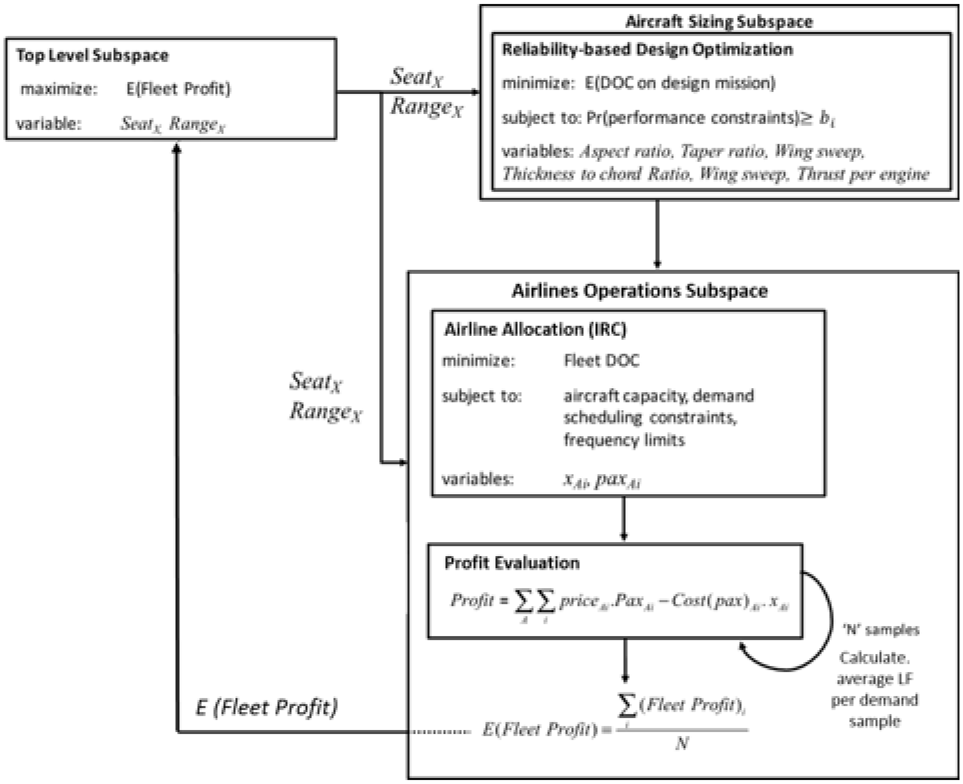


Figure 3. Subspace Decomposition Strategy for the Commercial Airline Application

Aircraft Sizing Subspace

This subspace is similar to the AMC work as described before. However, to accommodate different number of seats as required by the top-level problem formulation for the commercial applications, the sizing code needs to vary the size of the fuselage and the tail using an empirical relation established using the existing aircraft data. For this work, the uncertain parameters of choice, as appears below in Table 3, are selected based on subject matter expert opinion for illustrative purposes. A more formal approach of identifying most relevant factors would involve an Analysis of Variances (ANOVA) and a Design of Experiments (DOE) approach to identify the most statistically relevant design parameters influencing the aircraft design.

Table 3. Uncertain Parameters in the Commercial Aircraft Sizing Optimization Problem

Uncertain Parameters (ξ)	Lower Bound	Default	Upper Bound
C_{D_0} Multiplier [non-dim]	0.95	1	1.05
Oswald Efficiency Factor Multiplier [non-dim]	0.95	1	1.05
Thrust Specific Fuel Consumption Multiplier [non-dim]	0.95	1	1.05
Passenger Weight [lbs]	90	165	220

The RBDO formulation optimizes the expected performance metric of interest and ensures that the probability of satisfying the performance constraints is greater than or equal to the user-defined reliability level, considering the uncertainty present in this sub-problem. Here, we assume a triangular distribution for the uncertainties in each parameter; this will facilitate demonstration of the method, but better characterization of these distributions would improve the quality of the results. The aircraft sizing sub-problem includes performance constraints such as limits on takeoff and landing distances, second segment climb gradient, top of climb rate, and upper and lower bounds for the decision variables.

As mentioned earlier, at the solution of the RBDO problem, the resulting aircraft design has uncertain responses because of the input uncertainties (Table 3). Of interest for the airline operations subspace—the cost to fly the new aircraft on any route, the block hours needed to fly any route, the maximum number of passengers that the aircraft can carry on each route, and the takeoff distance of the aircraft—all follow probabilistic distributions.

Airline Operations

This subspace mimics an airline's operational behavior. The Interval Robust Counterpart (IRC) formulation recognizes and obtains the performance characteristics of the uncertain aircraft for the nominal and worst-case values of the uncertain aircraft design parameters of Table 3. We use these performance data in our allocation formulation to minimize the airline's fleet-level direct operating cost, while satisfying maximum predicted passenger demand on the route network. Here, the maximum predicted passenger demand comes from historical data available from the Bureau of Transportation Statistics; this provides a credible demand distribution for the problem, as if this historical demand were actually a prediction of future demand. Solving the allocation solution represents setting the airline's schedule, and then the approach samples the uncertain passenger demand that would fly on the set schedule and evaluates an expected profit considering the uncertain demand. To further capture seasonal variation in passenger demand, we set four different quarterly allocations. The purpose of considering each quarter's worth of data is to capture better the impact that seasonal fluctuations will have on the observed maximum number of passengers traveling on each route for a representative travel day. Average profit (or the expected profit) over all sampled demand for all the quarters then returned to the top level and appears as the top-level objective function.

Summary of Subspace Decomposition Approach in USAF AMC vs. Commercial Airline Applications

The main difference between the use of the subspace decomposition approach to the AMC and commercial airline cases are dictated by the nature of the payload for each aircraft type (pallets vs. passengers) and the statistical nature of the demand for transport (uncertainty, unstructured cargo vs. scheduled commercial flights). The details of differences in subspace modelling in both cases are summarized in Table 4.

Table 4. Differences in Subspace Formulations Between AMC and Commercial

	Subspace Decomposition Approach Application	
	USAF AMC Military Cargo	**Commercial Airline**
Subspace Level		
Top-Level	Requirements are number of pallets and range of aircraft. Use of Global Optimizer (NOMAD) to search design space	Top level requirements are number of seats and range of aircraft Perform Pseudo enumeration to search the design space
Aircraft Sizing	Fuselage sizing rules based on number of standardized pallets	Fuselage sizing rules based on number of seats
Fleet Operations	USAF AMC flight operations based on "as needed" basis for demand for cargo transport IRC formulation minimizes fuel consumption and enforces constraint on productivity. External demand sampling loop. Single IRC solution for each sampled demand set. Average fuel consumption of sampling returned to top level problem Use of aircraft assignment that tracks tail numbers of aircraft Demand sampling done by random sampling of starting locations for aircraft	Fleet operations based on BTS (BTS, 2015) data to model future prediction of demand (assumes demand in symmetric) IRC formulation minimizes operating cost while meeting the ever recorded maximum demand on a route for travel BTS data on historical airline data used to predict future demand distribution Scheduling done to meet maximum demand on all routes at same time Profit calculated through statistical sampling schemes on demand, and, includes ticket pricing model

Representative Results and a Posteriori Analysis

Military Air Cargo

Figure 4 shows the results from the multi-objective analyses of the 25-base network problem, using the subspace decomposition approach for the AMC case study (refer to

Figure 2). The plot shows the normalized expected values of the fleet-level metrics. Using normalized fleet-level responses help to identify the trends, and help to show the relative variations in fleet-level responses for different solutions to the multi-objective optimization problem. The fleet-level responses have been normalized with respect to the lowest expected values from the results of the scenario labeled “Fleet with five new A/C.” Each point in the “Fleet with five new A/C” scenario describes the optimal design of the new aircraft required to meet the specific fleet-level objectives. These results show the collection of optimal aircraft designs that would meet the fleet’s operational needs at each level of permitted fuel consumption or at each level of required fleet-wide productivity.

For three different solutions from the “Fleet with five new A/C” results, Figure 4 contains callout boxes that describe the values of the new aircraft requirement decision variables along with the values of the aircraft design variables. The trends in the fleet-level responses are as expected, with fuel consumption increasing as productivity increases. There appears to be a trend in the size of the optimal aircraft along the Pareto frontier for increasing productivity/fuel consumption values. For a normalized expected productivity and normalized expected fuel consumption value of 1.0, the optimal requirement decision variables of the new aircraft X are at the lower bounds for pallet capacity (16) and design range (3800 nmi). Moving from this point on the tradeoff plot towards solutions with increasing fleet-level productivity, the results suggest that larger pallet capacities for the new aircraft X can best meet the fleet-level objectives. There is not substantial evidence to determine whether these trends would generalize to other route networks or other similar design problems; however, the behavior is not unexpected, because the aircraft pallet capacity strongly drives the fleet-level productivity metric. Though it is intuitive that a larger aircraft would increase productivity, the optimal design features of the new aircraft X, such as the aspect ratio (AR_x), the wing loading $((W/S)_x)$, the thrust-to-weight ratio $((T/W)_x)$, and so forth, are reflective of the specific existing fleet and demand characteristics of the service network. For each solution in the plot, the assignments of the fleet of aircraft to routes are different to meet the actual demands better. The introduction of the five new aircraft (of type X) results in fleet-level fuel savings between 2.79% and 6.48% for the same normalized expected fleet productivity values, when compared to the case where only the existing fleet operates in the network.

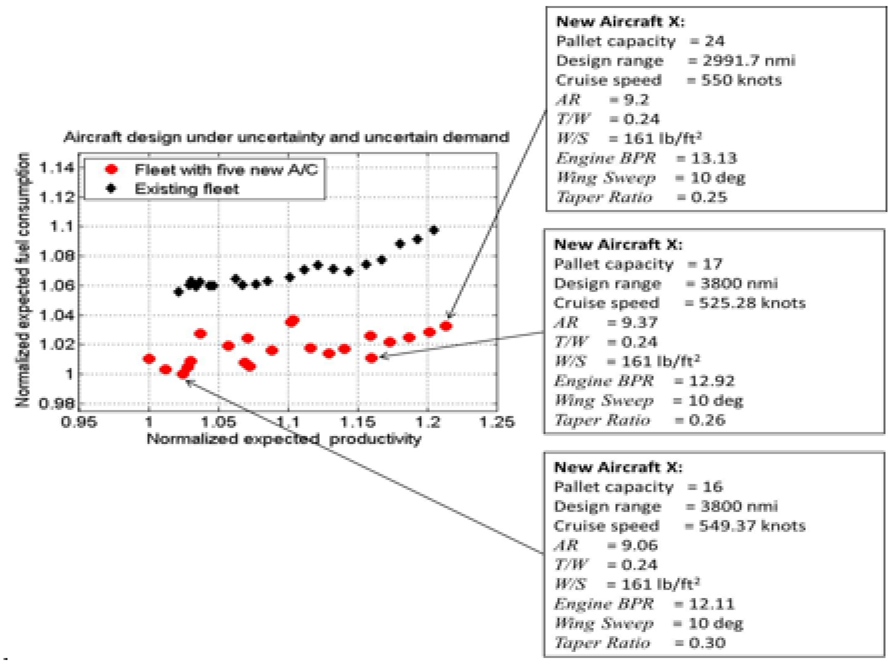


Figure 4. Results From Multi-Objective Analyses of 25-Base Network Problem

The solutions to multi-objective analyses present a way to perform “fuel/cost as an independent variable” type of trade-space analysis; this might be more obvious by switching the axes in the plot from Figure 4. These types of plots can help decision-makers/acquisition planners to analyze the trade-space and select the optimal requirements and design of the new aircraft that would achieve the desired level of fleet fuel consumption and productivity. For instance, a decision-maker can determine the level of fleet productivity available for a specific level of fleet fuel consumption; this fleet-level productivity value can then be translated to a specific (or bounded) level for the mobility airlift requirements that are set by the DoD in terms of tonnage of cargo transported per day. Having established the goals for the fleet-level productivity and fuel consumption, the collection of optimal aircraft designs required to achieve these fleet-level goals can be determined from plots such as those shown in Figure 4.

Posterior Analysis

Figure 5 shows the results from a posteriori analysis (200 samples) of a few solutions from the multi-objective analyses of the 25-base network problem. The dispersion in fleet-level fuel consumption does not show any discernible trend. However, the degree of dispersion in fleet-level productivity appears to decrease for solutions with increasing fleet productivity and fuel consumption values.

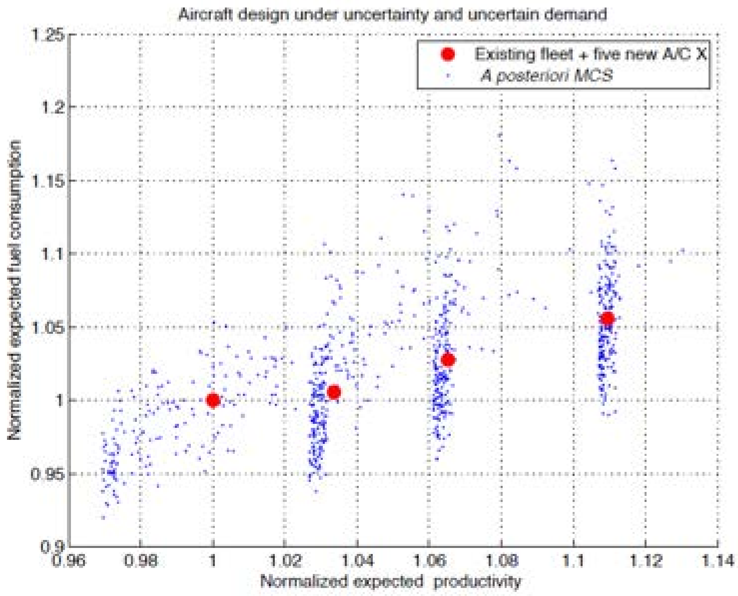


Figure 5. A Posterior Analysis for 25-Base Problem

Solutions with higher normalized fleet fuel consumption, in Figure 5, are more "robust" (less variance) in terms of fleet productivity. A possible explanation for this behavior is because the multi-objective analyses (using the e-constraint formulation) vary the limit value of the fleet productivity constraint, while minimizing fleet-level fuel consumption. If solutions that are more "robust" (less variance) to fuel consumption are desired, then the multi-objective analyses should vary the limit on the fleet-level fuel consumption constraint, while maximizing fleet productivity.

Decision-makers/acquisition planners can use such results to perform comprehensive exploratory analysis of the design space and identify regions in this design space that present significant viable or opportunities to reduce the fleet fuel consumption. For instance, AMC may need to incur "switching costs" (additional cost for training, maintenance and infrastructure due to the addition of a new aircraft type into the fleet) of integrating a new aircraft type into the fleet for relatively small decrease in fuel burn; however, the trade-space analysis (Figure 5 can help identify promising designs and inflection points, if they exist, where the decision to acquire a new aircraft type could provide significant benefits.

Commercial Airline

In the case of the commercial airline application problem, we solve a 31-route representative airline network as appears below. This network resembles a portion of the Northwest Airlines network before the merger with Delta and has its hub at Memphis.

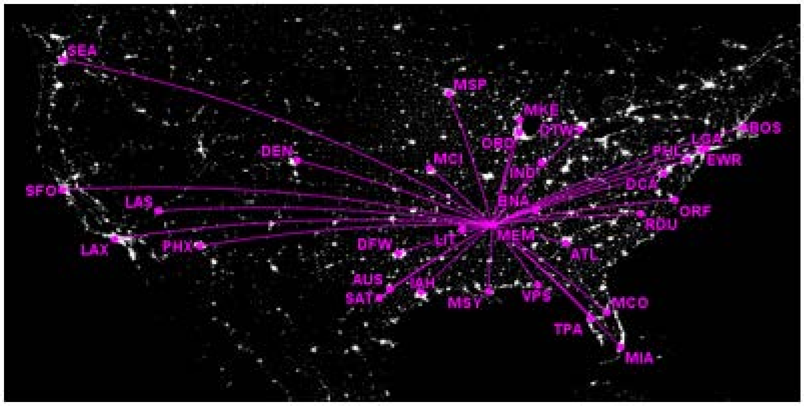


Figure 6. A 31-Route Network of the Example Airline Problem

The representative airline has the following fleet composition (Figure 7) and seeks to include five new yet-to-be-designed aircraft (from the aircraft sizing sub-space).

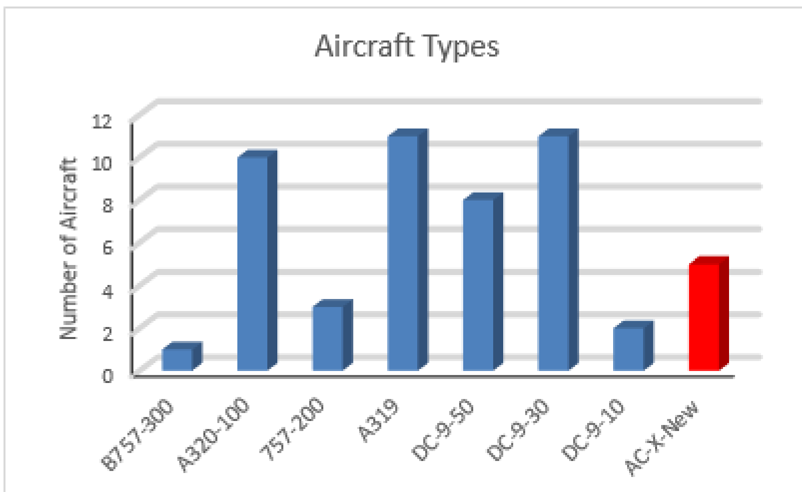


Figure 7. Fleet Composition of the Representative Airline

In this conceptual study, we used a pseudo-enumeration approach to address the top-level problem that uses the following range of discrete choices, as shown in Table 5. The interval values within the range for each of the variables is selected to more rapidly generate reasonable solutions at this stage of development in our approach—refinements in the grid space for the top-level enumeration scheme can be selected as required for more realistic problems.

Table 5. Design Variable Values of Top Level Problem for Enumeration

Range [nmi]	Seat Capacity
500	75
1200	150
1900	250
2600	

For each combination of design variables (4 range variables × 3 seat capacity variables = 12 enumerations points), we execute the overall subspace decomposition methodology shown in Figure 3. Figure 8 shows the profit data for all possible combinations of the enumerated top-level design variables from Table 5.

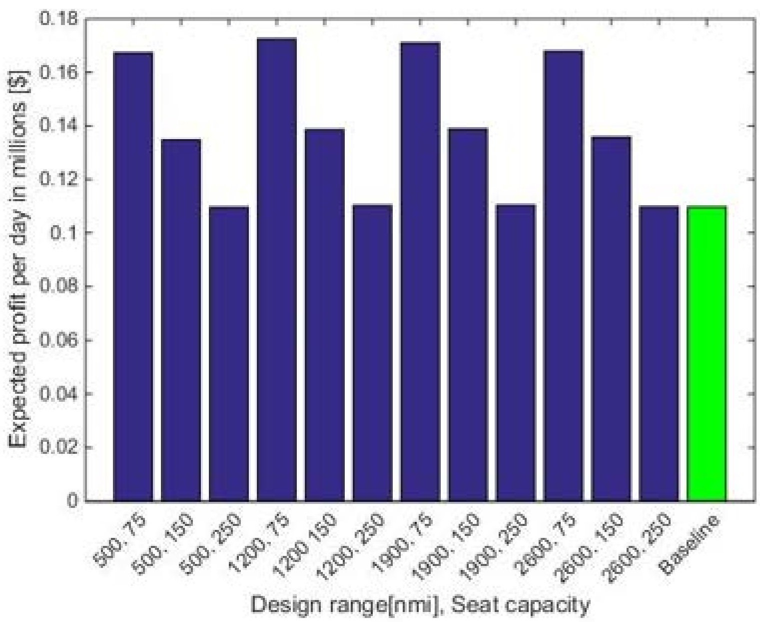


Figure 8. Expected Fleet Profit Values for the Combination ("Test Cases") of the Top-Level Design Variables (Green Denotes Baseline Fleet With No New Aircraft Type X Use)

The results show that the optimal seating capacity is 75 seats for the new aircraft, because the new aircraft is allocated on routes with average passenger demand of less than 110 passengers. Also, because the route distances of these routes in which the new aircraft is allocated are less than 1000 nmi (the longest route in the network is 1626 nmi), the optimal design range of the new aircraft corresponds to a distance of 1200 nmi. Further physical details of the optimal aircraft are retrieved from the aircraft design subspace problem that corresponds to the optimal range and passenger capacity values [1200nm, 75seats] and appear in Table 6.

Table 6. Optimal Aircraft Design

Optimal Aircraft Design Variables	
Aspect Ratio	12
Taper Ratio	0.3
Thickness-to-chord ratio	0.095
Wing Sweep [deg]	10.43
Wing Area [sq.ft]	664.76
Thrust per Engine [lbs]	9351
Passenger Capacity	75
Range [nmi]	1200

Figure 9 shows the utilization of each aircraft type in the fleet, over each quarter. In these plots, we note that most flights of the new aircraft design are allocated around the 500nmi range to fill in the travel needs. Given the number of aircraft available for each aircraft type, it is desired (as seen from the allocation results) to have a 1200nmi range aircraft, as it provides the option to be used on fewer long-range routes.

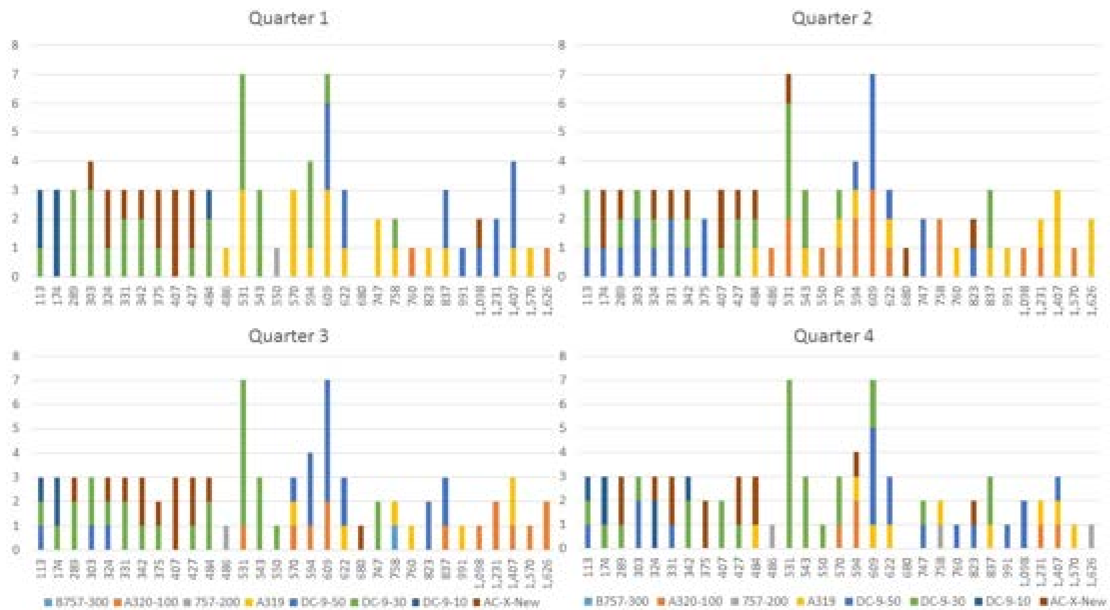


Figure 9. Distribution of Fleet Allocation Over Route for Each Quarter

Posterior Analysis

To validate the application of our framework, we performed a posterior analysis with a different set of 1000 samples. To generate this set of 1000 samples, we pick one sample for each uncertain parameter in the aircraft design subspace and performed an off-design mission analysis across all the routes in the network, keeping the aircraft design variables fixed to values obtain from the RBDO formulation. We then evaluate the performance characteristics of the aircraft and determine how many occasions these performance constraints are satisfied. Figure 10 below shows out of these 1000 samples how many times the aircraft performance constraints are met. Take-off distance seems to violate the most, as 78 of the 1000 samples did not meet the take-off distance criteria. The take-away from this plot is all the constraints are satisfied well within the 10%, which is our tolerance settings in the RBDO formulation at the time of designing the aircraft.

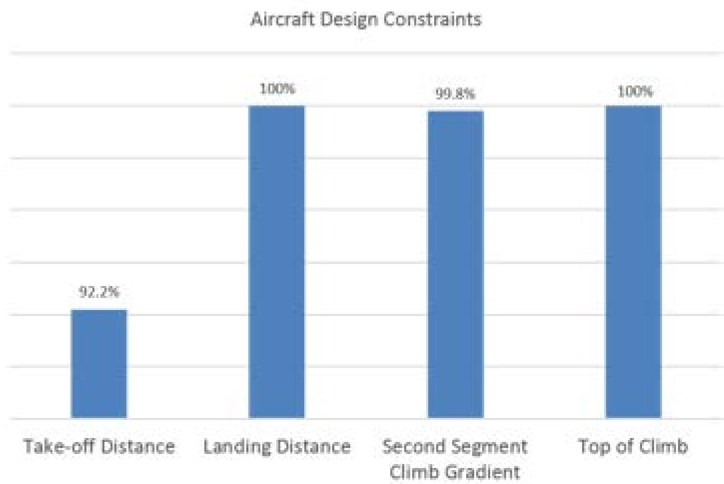


Figure 10. Percentage Satisfaction of the Aircraft Performance Constraints in Posterior-Analysis

Similarly, we performed a posterior analysis of the expected profit calculation, by sampling one instance of demand for every route and appears below in Figure 11. This demand sample combined with the extrinsic sample of the aircraft design subspace, both drawn independently, constitutes one sample for the posterior analysis. We repeat this step 1,000 times. Intuitively, one can say the expected profit from the posterior analysis should be around the same value as the original RBDO-IRC formulation run, if both of these methods handling the associated uncertainties well. This is confirmed in the plot below. We feel confident of our framework to address this type of problems, as attested via posterior analysis with 1000 independent samples.

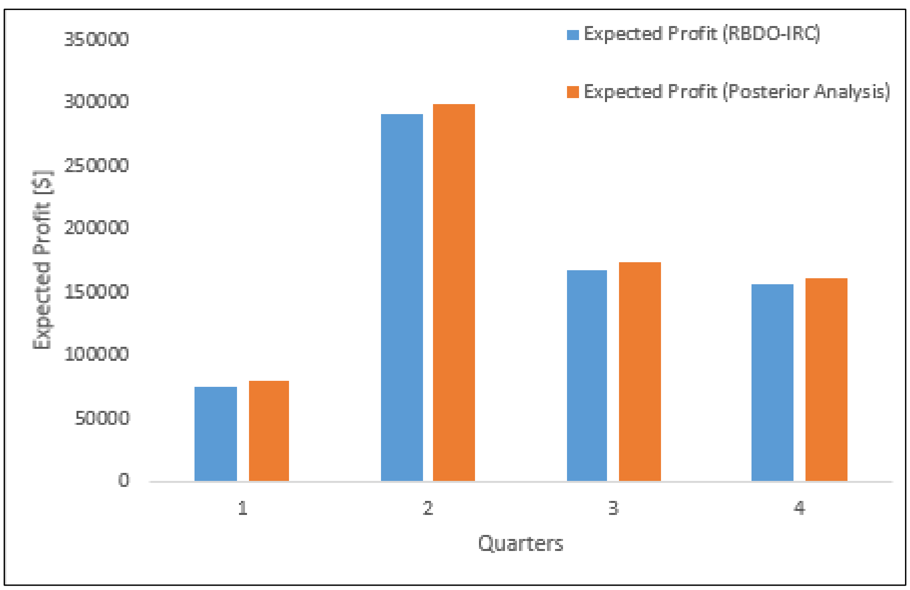


Figure 11. Expected Profit Comparison [Posterior Analysis]

Conclusions and Recommendations

In this paper, we have presented application of a subspace decomposition approach that better enables identification of design requirements of a new, yet-to-be introduced system (here, aircraft) towards improving fleet-wide performance metrics. The approach explicitly accounts for the impact that the new system will have on fleet-wide performance when used alongside existing systems within a fleet and also accounts for various data uncertainty that manifest in the problem. We have presented an application of the approach for commercial airline and military cargo airlift cases, demonstrating domain agnosticism of the approach. The approach is envisioned to be useful to relevant decision-makers within the general acquisition community (government, military, commercial) by enabling trade-off analyses between performance metrics of interest, and, under conditions of data uncertainty, thereby enabling a framework for robust decision-making on setting design requirements of a new, yet-to-be introduced system. Future work may encompass an extension of the approach to include additional relevant forms of domain-driven data uncertainty and further improvements in computational efficiency.

References

Brooke, A., Kendrick, D., Meeraus, A., Raman, R., & America, U. (1998*). The general algebraic modeling system, A user's guide*. Washington, DC: GAMS Development Co.

Bureau of Transportation Statistics (BTS). (2015). T-100 segment data.

Gertler, J. (2010). *Air Force C-17 aircraft procurement: Background and issues for Congress*. Collingdale, PA: Diane.

Govindaraju, P., & Crossley, W. A. (2015). Concurrent aircraft design and trip assignment under both design parameter and demand uncertainty. In *15th AIAA Aviation Technology, Integration, and Operations Conference* (ATIO).

Govindaraju, P., Davendralingam, N., & Crossley, W. (2015). A framework to determine new system requirements under design parameter and demand uncertainties. In *Proceedings of the 11th Annual Acquisition Research Symposium*. Monterey, CA: Naval Postgraduate School.

Graham, D., Bracken, J., Dalfonzo, J., Fedorochko, W., & Hilton, R. (2003). *Sustaining the Civil Reserve Air Fleet (CRAF) program* (No. IDA/HQ-P-3819). Alexandria, VA: Institute for Defense Analyses.

Greer, W. L. (2010). *An application of cost-effectiveness analysis in a major defense acquisition program: The decision by the US Department of Defense to retain the C-17 transport aircraft* (No. IDA-D-4218). Alexandria, VA: Institute for Defense Analyses.

Kendall, F. (2014). *Better Buying Power 3.0*. Washington, DC: Office of the USD(AT&L).

Lin, X., Janak, S. L., & Floudas, C. A. (2004). A new robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. *Computers & Chemical Engineering, 28*(6), 1069–1085.

Mane, M., Crossley, W. A., & Nusawardhana, A. (2007). System-of-systems inspired aircraft sizing and airline resource assignment via decomposition. *Journal of Aircraft, 44*(4), 1222–1235.

Montgomery, D. C. (2008). *Design and analysis of experiments.* John Wiley & Sons.

Roy, S., Crossley, W.A., Davendralingam, N., & Govindaraju, P. (2017). *Aircraft design optimization for commercial air travel under multi-domain uncertainties*. SciTech, 57th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference (2017), Grapevine, TX, January 2017.

Saliby, E. (1990). Descriptive sampling: A better approach to Monte Carlo simulation. *Journal of the Operational Research Society*, 1133–1142.

Taylor, C., & Weck, O. D. (2007). Coupled vehicle design and network flow optimization for air transportation systems. *Journal of Aircraft, 44*(5), 1479–1486.

Realistic Acquisition Schedule Estimates: A Follow-On Inquiry

Raymond Franck, Brig Gen, USAF (Ret.)—PhD, retired from the faculty of the Graduate School of Business & Public Policy (GSBPP), Naval Postgraduate School (NPS) in 2012. He retired from the Air Force in 2000 in the grade of Brigadier General. His active duty career included a number of operational tours and staff positions, and head of the Department of Economics and Geography, USAF Academy. His published work includes many journal articles and research reports in military innovation and defense acquisition management. [cfranck215@aol.com]

Gregory Hildebrandt—PhD, has had an Air Force career including assignments as an acquisition officer and Air Force Academy faculty member, plus assignments at the Central Intelligence Agency and Office of the Secretary of Defense. Following his Air Force retirement, he has continued service with the RAND Corporation and NPS. His published work includes a number of journal articles in defense economics and RAND reports on acquisition issues. [ggh324@gmail.com]

Charles Pickar—DBA, is a member of the NPS faculty where he teaches project management, defense acquisition, and systems engineering. Before joining NPS, he led the Applied Systems Engineering Program Area in the National Security Analysis Department of the Johns Hopkins University Applied Physics Laboratory. He is a retired Army officer with extensive experience in the U.S. defense industry, to include director and VP levels at Lockheed Martin, Northrop Grumman, and SAIC. He is the current Chair of the Systems Education Technical Committee of the IEEE Systems Council. His research and published work focuses on applying systems engineering and system dynamics analytical approaches to defense acquisition problems. [ckpickar@nps.edu]

Bernard Udis—PhD, is a Professor Emeritus of Economics at the University of Colorado at Boulder. He has also served at the U.S. Air Force Academy and the U.S. Arms Control & Disarmament Agency. His NATO research fellowship examined the costs and benefits of offsets in defense trade. A recognized authority on the economics of defense, his published work includes three books, plus a number of book chapters and journal articles. [bernard.udis@colorado.edu]

Abstract

In a 2016 proceedings paper for this symposium (Franck, Hildebrandt, & Udis, 2016), we outlined and discussed a research agenda with an aim of more realistic acquisition program scheduling estimates, especially for the development (SSD) phase. This paper is intended to continue pursuit of that agenda, with the aim of demonstrating its feasibility and discussing methods of analysis. Accordingly, this paper is presented in four parts: the promise of Systems Dynamics as a schedule-estimating paradigm; a preliminary case study of cost-performance-schedule tradeoffs in the F-35 program; the development of measures of effectiveness for contemporary air-to-air combat; and finally an examination of data sources and empirical models for program and contract schedule uncertainty.

Introduction

Among other things, our proceedings paper for the 13th Annual Acquisition Research Symposium proposed a research agenda intended to enable more realistic acquisition schedule estimates (Franck et al., 2016). This paper pursues that agenda along a number of lines—with the aim of finding or exercising methodologies that can be applied to developing more realistic schedule estimates.

Program schedule time could possibly be analyzed and forecast according to the following menu (p. 99):

- An orderly function involving key variables. This would lead to “schedule estimating relationships.”

- A result of a series of management decisions intended to produce the best program. At the macro level, this is a set of cost-performance-schedule tradeoffs, which are related in complex, imperfectly understood ways.
- An outcome result arising from the interactions among a set of tasks needed to complete the program. Among other things, this raises the question of tracing through the sometimes tangled relationships among various parts of the program (perhaps unplanned).

In that vein, we've undertaken to explore methods to better understand all three approaches.

We explore ways to identify the key variables for estimating schedule length—with a view to developing "schedule estimating relationships." In the first section, Pickar offers an overview of Systems Dynamics methodology in understanding the evolution of acquisition programs over time. Ability to model events and processes has potential to provide insights into the effects of decisions and unplanned events on acquisition timelines.

Another way of better understanding those decisions, unplanned events, and their outcomes is through case studies. In the next section, Franck and Udis undertake a preliminary case study of the F-35—with a focus on system tradeoffs.

In the section titled Better Understanding Acquisition Schedules Through Better Performance Measures, Franck assumes that system performance is a promising variable for schedule estimating relationships, and essays a preliminary approach to performance measurements for contemporary air combat systems—beginning with variants of Lanchester models of combat. This section also suggests that improved methods of performance quantification can shed light on cost-performance-schedule tradeoffs (the macro-level tasks of program management).

Finally, in the Program and Contract Schedule Uncertainty section, Hildebrandt undertakes continued study of schedules, and variances thereto, using statistical and econometric approaches with particular emphasis on data sources.

Schedule Estimating Methodologies: The Promise of System Dynamics

> Most resource-constrained project scheduling research efforts have been made under the assumption that the project scheduling world is deterministic, while uncertainties during project execution are quite common. (Herroelen, 2005)

We continue this discussion of scheduling by comparing the traditional schedule estimation process with system dynamics. Clearly DoD project managers, engineers, estimators, and contracting officers have significant experience in developing project schedules. They also use the most modern scheduling software and processes. These tools are mainly based on the critical path method (CPM) and the program evaluation and review technique (PERT), which may be insufficient for the scheduling tasks at hand.

Accordingly, we use systems engineering methods to reduce the complex to its components rendered as the Work Breakdown Structure (WBS), and further reduce work to the work package level. These work packages are then defined and resourced and become the basis for not only scheduling, but cost and risk as well. WBS provides a decomposition of the project to a level that provides visibility of the work, as well as work progress. Once the effort is defined, schedule and the accompanying cost can be identified. The estimating effort is closely related to cost estimation and uses many of those same methods including analogy, parametric, algorithmic models, and expert judgment. Of these, expert judgment should be able to estimate the effort necessary to schedule the development tasks (Hughes, 1996). In fact, expert judgement is a first step for most cost and schedule estimating (Abdel-Hamid & Madnick, 1983).

Current Scheduling Methods

Scheduling tools (mainly based on the CPM and the PERT) apply a network approach to define critical activities, slack, and the overall time required to complete the development. The network approach also provides the basis for cost estimation, resource allocation, management focus and risk assessment, and provides a visual flow of the effort. However, given their work package or task level, CPM/ PERT force managers tend to develop a myopic view—for good reasons. The disadvantage is less focus on a systems view of the development effort. This leads to two major difficulties. First, CPM and PERT take a static view of project activities (Balaji & James, 2005), which fails to account for the relationships and interdependencies inherent in complex projects. Second, is the basic assumption that work proceeds as planned in the network—that there is a direct flow from work to be done, to work accomplished. That is, every CPM task has a discrete start and end—work is either started or not, finished or not. There is no accommodation for work that might not be done correctly or to the required quality (Cooper, 1993c).

Reality and empirical research show that work scheduled is not always completed to the necessary quality and therefore must be redone (Cooper, 1993b; Cooper, Lyneis, & Bryant, 2002; Rodrigues & Bowers, 1996).

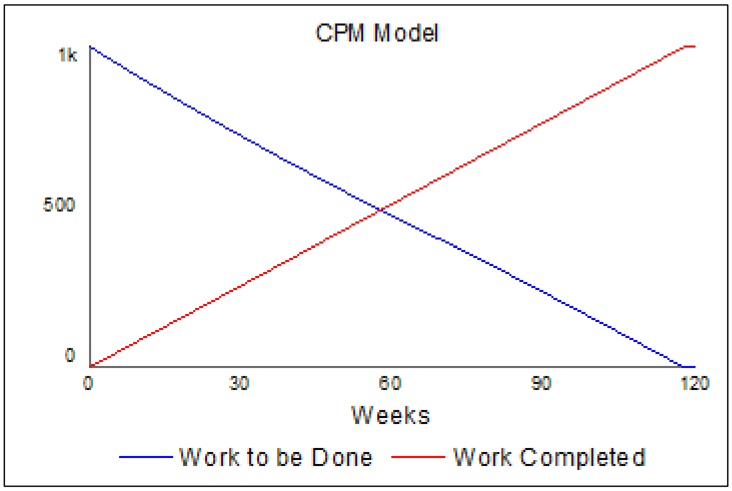


Figure 1. Generic Project Without Rework

Figure 1 shows a simplified, generic project with 1,000 tasks, executed by 10 people at approximately 90% productivity rate. The X-axis is weeks, and the Y-axis shows the number of tasks. The graph shows both a steady reduction in work to be done, and an equally steady increase in the work completed. The graph shows completion of these 1,000 tasks at week 117.5. The graph represents a deterministic view of scheduling that doesn't account for delays or changes in complex projects, whatever the cause. Proponents of CPM and PERT recognize the limitations resulting from the deterministic approach and have made attempts to incorporate more realism by adding probability measures to the estimated times, but the root problem remains (Kerzner, 2013; Moder, Phillips, & Davis, 1983).

The Dynamics of Projects

System Dynamics, a relatively new field, was developed by Forrester in the 1960s. It provides a conceptual modeling and simulation tool that uses differential equations to track the interdependencies, flow and dynamics of a process. Initially advanced as a management tool, it is also used to address broad policy issues, as well as project management (Forrester, 1987, 1995; Sterman, 2000; Williams, 2002). System dynamics “deals with the time-dependent behavior of managed systems with the aim of describing the system and understanding through quantitative and qualitative models” (Coyle, 1996, p. 5). It provides the project manager a different, system-level view of the schedule and its execution—to include cause and effect relationships; non-linearity; and understanding of the impact of feedback loops (Godlewski, Lee, & Cooper, 2012). System dynamics offers the possibility of augmenting traditional scheduling methods to provide not only better scheduling estimates, but better visibility of project status as well. The field gained recognition as a contract dispute resolution tool in the 1970s when it was successfully used to argue a weapons systems contract dispute with the U.S. Navy (Cooper, 1980).

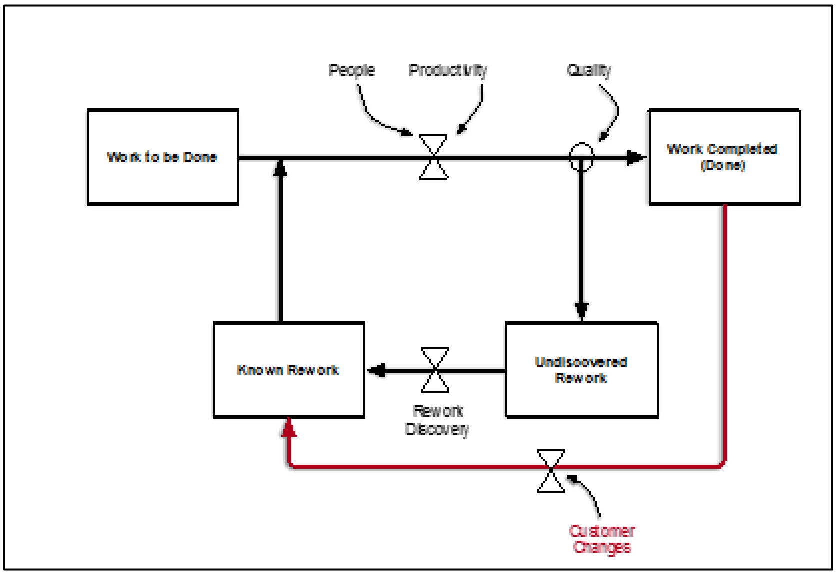


Figure 2. The Rework Cycle

The rework cycle is a fundamental system dynamic concept first articulated by Cooper (Cooper, 1993a, 1993b, 1993c). Figure 2 depicts the rework cycle. Unlike the CPM tracking of discrete tasks system dynamics monitors flows, the basic flow of work in a development is from *work to be done* (tasks or work packages) to *work completed*. Connecting that flow is a “valve” that regulates the flow. In the rework cycle, that flow is determined by *people* (numbers, skills, availability) and *productivity*. People times productivity provides a flow rate, for example, tasks per week. *Quality* is another modulator of the flow of work. Quality is simply a measure of whether the task was accomplished correctly and completely. Given the exploratory nature of research and development efforts, it is entirely possible that a planned development task fails to accomplish the task goal, and the task must be redone. Similarly, people may be operating at a high level of productivity, but not producing quality work.

There are two types rework, known and undiscovered. These categories are integral to the nature of weapons system development. Developmental test does identify some of the work that needs to be redone, and that work flows to the known rework stock. However, there is work that may pass developmental test, but is later found to be deficient (software “bugs” are a good example). Those deficiencies may not be discovered for significant amounts of time. Those deficiencies may also cause follow-on developmental efforts to slow or fail until they are finally discovered. Rework is a known issue for experienced project managers. Understanding the impact of the rework cycle coupled with the CPM network can provide a tool to develop better schedule estimates.

The red arrow (lower right) in Figure 2 is an example of the dynamic, causal effects system dynamics can track. In this case, customer changes add to the basic rework cycle by measuring the impact customer changes can have on a development. In the DoD, "customer" includes the program office, the acquisition chain of command, and the requirements community. Customer changes are shown as a valve that affects the flow of work that may be complete, by delaying the flow of the work both from changes as well as delays from information gathering, preparing, and reporting. For instance, the GAO found that the F-22 Increment 3.2B Modernization spent 3800 staff days to prepare 33 milestone documents and present 74 briefings for the Milestone B process (GAO, 2015). This work had a cost of some $10 million. These 3800 staff days obviously would also have impacts on the schedule, potentially more significant than on cost. Another schedule impact is driven by the customer changing requirements, specs, or simply delaying responding to information guidance requests from the PMO.

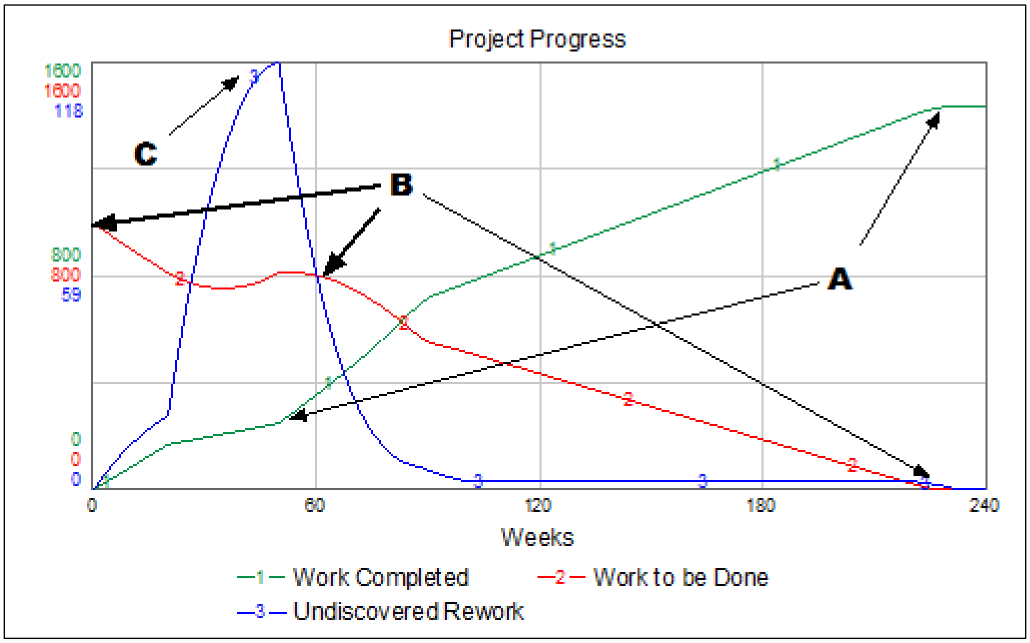


Figure 3. Effect of Rework on Generic Project

Figure 3 shows the results from the same generic model used in Figure 1, but this time incorporating the impact of the rework cycle. The X-axis shows time, and the Y-axis indicates number of tasks. Line A, shows the Work Completed, line B shows Work to be Done, and line C shows a generic calculation of Rework. Comparing line A in this graph to that of work completed plot in Figure 1 demonstrates the effects of rework. In this case rework peaks at week 48 (line C), and is estimated at 75%. This means three of every four tasks must be redone, a conservative estimate by some measures especially when considering software development projects (Cooper & Mullen, 1993). Similarly, line B (Work to be Done) shows a much longer completion time than that of Figure 1. Completion time in this model run is 229 weeks, an increase of 111.5 weeks over the generic model in Figure 1, an almost 100% increase in schedule. Another way of considering the impact of rework is that instead of the 1,000 tasks originally required, the number of tasks completed was 1,437—a significant increase in work requiring more time and money.

While this is an elementary model, it demonstrates that something as simple as rework (because of quality) can have a significant effect on project schedules. It also validates an oft repeated axiom in defense acquisition—a design freeze, made possible by effective systems engineering "portends better program outcomes" (GAO, 2016b).

Summary

Weapons system development is itself a system—a collection of inputs and outputs, events and activities that interact over time to produce a unified whole. Dynamics describes the forces or properties that drive change within a system over time. Thus, system dynamics describes the change within a system over time—a stochastic perspective. And managing change is key to effective program management.

The deterministic schedule estimating methodology in use in the DoD today is defined by CPM. This perspective assumes each task is discrete, and does not account for interaction and interdependency. It also assumes each task is either completed correctly and completely the first time, or that any necessary further work is included in downstream estimates (Reichelt & Lyneis, 1999). DoD project managers know this is not the case. Thus, system dynamics combined with current CPM-based approaches provide the potential to meet the need to better estimate weapon system development schedules.

The F-35 Case: A First Look at Tradeoffs Among Performance, Cost, and Schedule

A useful step toward better schedule estimation is to better understand the nature of the major tradeoffs in acquisition program management: cost, performance, and schedule. Application of paradigms such as Systems Dynamics (above) is one such approach. Another is case studies of actual acquisition programs.

Accordingly, this section provides a preliminary case study of tradeoffs in the Joint Strike Fighter (F-35) program. We pose three questions:

1. What sorts of trades were made in the F-35 program after the 2001 source selection decision?
2. What were the consequences of those trades?
3. How and why were those trades made?

JSF Intentions at the Start

During the early 1990s, a number of tactical fighter programs and initiatives were cancelled—with the common theme being affordability. These included the Navy's A-12 attack fighter, the Naval Advanced Tactical Fighter (F-22 variant), and the Air Force Multi-Role Fighter (DoD, n.d.-c). The focus shifted to jointly designed and procured systems. This included a number of STOVL-capable initiatives, to include the CALF (Common Affordable Lightweight Fighter, 1993–94)—intended to develop "technologies and concepts" to support Harrier replacements for the USMC and Royal Navy, plus a highly-common conventional aircraft for the USAF (DoD, n.d.-c).

The surviving initiatives were combined under the JAST (Joint Advanced Strike Technology) program, which began operations in January of 1994. JAST's original charter was to "mature technologies, develop requirements, and demonstrate concepts for *affordable* next-generation joint strike warfare" (emphasis added; DoD, n.d.-b).

This theme was strongly echoed in a Defense Science Board (DSB) study published later that same year (DSB, 1994). The DSB's findings and recommendations strongly emphasized (a) well-specified operational capabilities (esp. p. ES-2), (b) "affordable processes and end products" (ES-2), (c) technical conservatism ("a customer for technology, not a … developer" [ES-2]), and a strong emphasis on affordability (calling for "revolutionary improvements in affordability" [ES-4], described as the "key enabling technology" [ES-5]).

Similar themes appeared in a 1997 paper by RADM Steidle. He emphasized "doing business differently" through "streamlined, nontraditional business approaches" (p. 7). These included Government-Industry teamwork, Cost as Independent Variable (CAIV, in the interests of affordability), and using best practices ("sound ideas for improvement"; Steidle, 1997, p. 7).

These practices included pursuing affordability through a high degree of commonality among all services' variants (Steidle, 1997, p. 9), choosing autonomic support as "the … logistics response to … enhanced onboard weapon system diagnostics" (p. 9), and pursuing extensive cost-performance tradeoff studies (pp. 9–11).

Accordingly, the first JSF Selected Acquisition Report (SAR) stated that "the cornerstone of the JSF Program is affordability" (Steidle, 1996, p. 2).[1] And, fully consistent with DSB recommendations, committed the program to "fully validated, *affordable* operational requirements," (emphasis added), technical risk reduction, and "demonstrating operational concepts" (p. 3).

What Resulted

> "There's not a more complex program on the planet. … We did it because we had this grand vision."—(then) Maj Gen Christopher Bogdan (2012)
>
> "Our complexity reach exceeds our engineering grasp."—Anonymous DoD official (2011)

It seems clear the JSF we originally wanted is not the JSF we're getting. The CALF (Common Affordable Lightweight Fighter) initiative was intended to design demonstrator aircraft to include affordability analyses—with few hard constraints. And the JAST (as noted above) heavily emphasized affordability and technological conservatism. The JSF is not even close to these original visions, as illustrated in Table 1.

The F-35 is a much more capable aircraft than the Joint Strike Fighter originally envisioned. According to Lockheed-Martin, the F-35 has (or will have) a wide range of capabilities, of which air-to-ground attack is merely one (LM, 2017). More importantly, the F-35 that is emerging embodies new concepts of warfare based on situational awareness shared through secure networks. The overall effect according to one source is that the F-35 "*will be part of a <u>strategic</u> <u>transformation</u>*. The ability of the aircraft working with the other elements … will allow tactical maneuver to have a strategic consequence" (Laird, 2012).

[1] This exact statement appeared in JSF SARs for more than a decade, and is still posted on the DoD's JSF History website (last accessed in 2017).

Table 1. Joint Strike Fighter vs. CALF

Common? No	"… three separate programs that have common avionics and a common engine"[a] (Bogdan, 2012); 20% common[b]
Affordable? Maybe.	Significant cost growth[c]
Lightweight? No.	Weight about the same as F-15
Fighter? Depends.	"No": glorified F-117[d]; clearly inferior to the Su-35[e]; "It doesn't matter": a game-changing situational awareness machine[f]

Notes: a. Bogdan (2012); b. Seligman (2017); c. GAO (2016a, p. 10); d. Anonymous AF Fighter Weapons School Instructor (2015); e. Airpower Australia (2017); f. Laird (2012)

Consequences: The Price of Performance

However, increased performance has come at a high price—in both schedule and cost. The multiple program redefinitions ("re-baselining") complicate measurement somewhat, but nonetheless cost has grown significantly, and schedule has stretched (GAO, 2016, p. 10). The indirect effects of F-35 funding additions and delayed operational capability have likewise been significant (Sweetman, 2012; Freedburg, 2017; Tirpak, 2017b).

The F-35 program has encountered a large and complex list of complex problems. Those difficulties, and the management actions to deal with them, similarly constitute a highly complex history. Highlights include the Quick Look Review of 2011 and continuing operational testing issues.

The Quick Look Review was an OSD-mandated study of F-35 development issues. The team found no show-stopping "fundamental design risks," but did identify a total of 13 issues, of varying significance. Program concurrency was identified as a matter of overarching concern (Axe, 2011).

Since then, both the Marine Corps and Air Force have declared Initial Operational Capability for their models, and a number of program issues have also been resolved. However, significant problems still remain. The F-35 section of the *FY2016 Annual Report* from the DoD's Director of Operational Test and Evaluation (DOT&E; 2016) identified a number of concerns including the following:

- Testing schedule delays (p. 47),
- Air frame issues including vertical tail attachment bushing fatigue (p. 48),
- Mission Data File production capacity (p. 49),
- Autonomic Logistics Information System (ALIS) capabilities for full combat capability expected after declared IOCs (p. 68), and
- Weapons release limitations (p. 62).

None of these are trivial. For example, "if we don't get ALIS right, we don't fly" (Bogdan, 2012). It's a safe bet that these problems, plus the others DOTE identified, are solvable. It's a safer bet that they will further delay, and make more costly, fielding of F-35 combat capabilities.

How and Why Were the Trades Made?

As a first step in answering this question, we drafted a list of keywords associated with information and network-based concepts of operation cited by current F-35 advocates. We then consulted readily available sources for consideration.

Unclassified documents such as the Joint Strike Fighter Selected Acquisition Reports (starting in 1996) unearthed surprisingly little congruence with current F-35 concepts of operation. For example, "situational awareness" discussions emphasized the JSF as a consumer of offboard information, with much less attention to the aircraft as a source (or sharer) of information. Similarly, the term "network" appears with a fair degree of regularity, but mostly referred to means of testing and evaluation—with some addressing the JSF's role in a future system-of-systems architecture. There was, for example, no mention of the battle-management role now widely discussed (e.g., Weisgerber, 2016).

Our inquiry did, however, yield some interesting bits. These included RADM Steidle's (1997) rationale for autonomic logistics as a logical extension of onboard fault-detection systems. In retrospect, this could have been the first along the primrose path to current ALIS difficulties. We also have Gen Bogdan's (2012) intriguing reference to a "grand scheme" for the JSF. We hope to learn more in future inquiries.

Answering the Questions

Were cost-performance-schedule trades made? Our answer is "yes." Relative to the JAST conception, there were major trades made for improved performance. This conclusion seems unsurprising, and is likely a matter of common knowledge among members of the acquisition community.

Their consequences? Choosing higher performance was a major factor in explaining cost overruns and schedule delays. Parts of Gen Bogdan's "grand vision" (2012)—such as the helmet mounted display unit and autonomic logistics information systems—have resulted in long-term difficulties, causing delays and costing money. In addition, attempts to keep the program closer to planned schedule appear to have increased program concurrency, which in turn has been a cause of cost growth and schedule delays.

How and why were the tradeoffs made? From the references consulted, we were unable to find more than fragments from the record of how and why the JSF went from a conservative, relatively specialized concept aircraft to an "extraordinary" multi-role combatant (Miller, 2016).

This has been an initial inquiry, and the answers appear beyond the limits of the open literature. More complete answers entail more work, to probably include field interviews.

Better Understanding Acquisition Schedules Through Better Performance Measures

There are a number of good reasons to estimate combat performance for modern systems (and systems of systems) that encompass both acquisition management and defense planning (Table 2 below). For example, the new fifth generation of combat aircraft conducts air warfare in significantly different ways. As one Air Force officer put it, "With fourth-generation fighter airframes, speed and energy equaled life and survivability. In the fifth-generation realm, information equals life" (Fraioli, 2016).

Hence, planning for information-age combat forces would benefit from better understanding of combat effectiveness. As a Marine fighter pilot put it, "We need to do a better job teaching the public how to assess a jet's capability in warfare" (Lockie, 2017).

Table 2. Air-Air Combat Tasks (Kill Chain)

TASKS	CAPABILITIES
Cueing	Intelligence, Surveillance, & Reconnaissance (ISR)
Detection, Identification, & Tracking	Sensors
Allocation of Forces	Command & Control (C2)
Engagement	Platforms & Weapons
Assessment	ISR

Second, The JCIDS (Joint Capability Integration and Development System) Instruction recommends "effective cost, performance, schedule and quantity trade-offs" as being highly conducive to successful acquisition programs (CJCS, 2015, p. A-9). Making those tradeoffs well informed presupposes there's some useful way to measure performance.

Finally, better understanding of performance has real potential for explaining (and predicting) system development schedules—analogous to use of performance indexes (scalars) in explaining cost of previous generations of fighter and attack aircraft (Hildebrandt & Sze, 1986). That is, better measures of performance may lead to better schedule estimating methods.

Hence, our purpose in this section is to start a discussion about measuring performance in Information Age warfare. We hope readers find some useful insights in what follows.

The basic Lanchester model is primarily about the engagement task. But military affairs have gotten more complicated. The variants discussed below are a first step in widening that analysis. This section considers "detection, identification & tracking" (shortened to "detection"), Allocation (C2), in addition to expected results of engagements.

Lanchester Aimed-Fire Model

Results in the basic Lanchester aimed-fire model[2] (best known of the genre) depend on numbers (B and R) and "lethality" (b and r). At any given moment in a Lanchester battle,

$$dR/dt = -b * B \text{ and } dB/dt = -r * R \quad (1)$$

Blue (Red) lethality, b(r) depend on Blue's (Red's) rate of fire, accuracy, and lethality of munitions. Blue (Red) lethality also depends on Red (Blue) ability to counter, evade, or nullify the effects of that fire.

[2] Studies in the Lanchester tradition have resulted in numerous, varied, and highly ingenious analyses. Taylor (1983) and Bracken, Kress, and Rosenthal (1995) are but a few examples.

While casualty rates (r*R and b*B) are important, casualties as fractions of the respective forces determine the winner. Thus,

$$(dR/R)/(dB/B) = (b * \left(\frac{B}{R}\right))/(r * (\frac{R}{B})) \; (b*B^2)/(r*R^2),$$

And numbers count more than unit lethality in determining military effectiveness. That is,

$$E(R) = r*R^2, \text{ and } E(B) = b*B^2. \qquad (2)$$

A Lanchester Model with Probability of Detection Varying

The basic Lanchester model embodies assumptions that do not reflect information-age air combat well. Among other things, we should explicitly consider the possibility of not all combatants present being detected well enough to support a targeting solution.

This affects casualties inflicted. If, for example, half the Blue force disappears (in effect) from Red's situational picture, then available Red units must concentrate their efforts against fewer targets—with Red's fire having less effect on the entire Blue force. Basically, we assume Red forces concentrate their fires on the detected Blue forces, and must leave the rest alone.

In general, therefore, we can describe rate of casualties for Red on Blue as follows:

If $P_{dR}*R>B$, then $dR(t) = -b*B(t)*$, (3)

and if $P_{dR}*R<B$, then $dR = -[(1-PM_R)\text{^}(B/(R*P_{dR})]$, (4)

is a reasonable approximation.

Hence, varied detection probabilities (degrees of stealth) can change the relationship of relative force sizes in the outcome. Also noteworthy is that stealth needn't be absolute to be operationally significant.

One summary of our simple exercise is the following estimating equation:

dR/R = - 0.22 + 0.30*(B/R) + 0.12*PdR +0.19*b + 0.04*C2 – 0.13*(PdR^2) – 0.02*(b^2) – 0.10*(C2^2) + 0.49*(PdR*b) + 0.08*(PdR*C2) +0.13*(b*C2) (5)

where the variables' are defined as follows:

(dR/R) is the estimated proportion of Red force attrited by Blue fire,

(B/R) is relative force size (Blue vs. Red) at the start of the engagement,

PdR is probability of detecting Red units (0 to 1),

b is Blue units' lethality vs. Red units (0 to 1),

C2 is degree of command and control (0 to 1).

The linear variables provide the "direct" effects. The squared terms indicate how the direct effects are modified as their size increases. If the coefficient is negative (as all are), then incremental effect decreases as the variable increases—also known as "diminishing returns."

The multiplicative terms indicate the interactions among the variables. If the coefficients are positive, then there is a "synergistic" effect among the variables. Also, it means that ability to substitute one attribute for another decreases as it increases—"diminishing marginal rate of substitution."

Figure 4 illustrates this point. If the probability of detecting Red (PdR) decreases from 0.8 to 0.6, an increase of 0.05 in Blue lethality (b) can compensate. However, it takes an increase of 0.25 in b to compensate for a reduction PdR from 0.4 to 0.2.

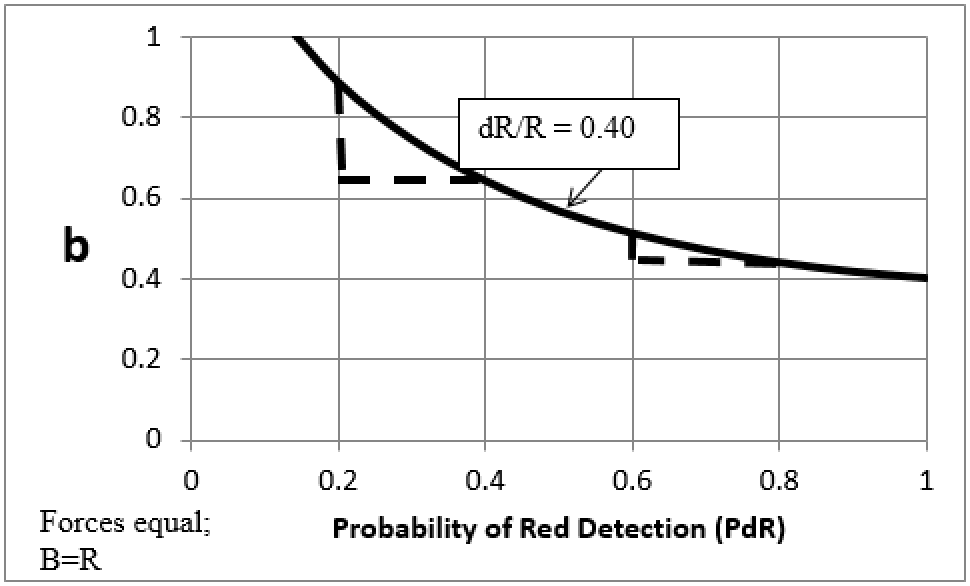


Figure 4. Red's Proportion Lost vs. Blue Lethality (b) and Probability of Red Detection (PdR)

Measuring Air-to-Air Firepower

What's above is a first step in measuring air-to-air firepower. Figure 5 provides an approach to a more complete measure of capability. Air combat power is basically ability to accomplish the stated tasks given. The goals hierarchy essayed in Figure 5 centers on the "engagement" task with "fire power" as metric of capability.

One should not expect all the sub-goals (like sensor vs. weapons capabilities) to be equally important. Likewise, we can expect interactions among the sub-goals in enhancing firepower (as shown in Equation 5 above).

The task ahead is to combine good analysis and expert judgement to formulate an estimate for "firepower" as a function of Command Control, Force Size, Sensors, Shared Situational Awareness and Weapons. While this is large and complicated effort, it also appears to be tractable.

Closing Comment: Fighting Outnumbered and Winning

"The pilots will … penetrate contested space and are likely to be outnumbered by adversary aircraft" (Miller, 2016).

However, the advantages of more stealth[3] are more completely observed in force-on-force models. The Air Force Chief of Staff speaking at a recent event dismissed one-on-one comparisons fighter types (such as F-22 vs. J-20) as not really relevant, and indicated that net assessments of network vs. network are crucial (Tirpak, 2017a).

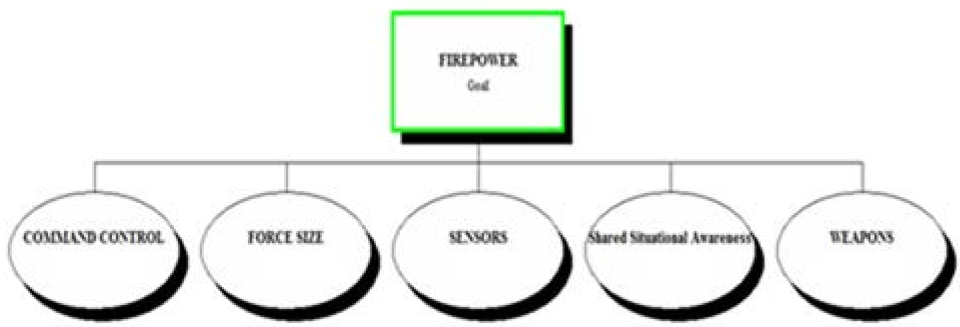


Figure 5. A Draft Goals Hierarchy for Estimating Air-to-Air Combat Potential

To illustrate, Figure 6 displays Blue stealth (PdB) needed to win vs. the probability of detecting Red units (PdR). Within the scope of a simple force-on-force model, Blue wins in all combinations of detection probabilities in the lower right; Red wins to the upper left. Toward the right side of the figure, weapons matter more; at the left, stealth matters more.

The horizontal scale (0–1) compared to the vertical (0–0.4) is worth noting. It takes a significant Blue stealth edge to overcome the advantage of superior numbers. Or as Air Combat Command's General "Hawk" Carlisle put it, "Fighter technology really isn't the problem. It's really about numbers"[4] (Tirpak. 2017c).

That said, however, increasing Blue lethality (b) also provides considerable advantage in countering superior numbers. These results are preliminary and suggestive. At best they provide some useful insights. Next steps are to finish development of this model further and then verify and validate it.

Such a model, combined with appropriate expert judgements, can lead to useful performance measures. The performance measures, in turn, can provide inputs for credible schedule estimates.

We have departed from measuring platform-on-platform capabilities and into the realm of force-on-force. In an era of systems of systems and network vs. network that seems appropriate.

---

[3] We mean "stealth" in a broad sense: the ability to avoid being engaged. Detection, identification and tracking sufficient to support an engagement can be denied by a number of means—to include various forms of information and electronic warfare, as well as through low-observable platforms.
[4] This statement may appear more strongly than Gen Carlisle intended it. Nonetheless it points out the continued relevance of comparative force sizes.

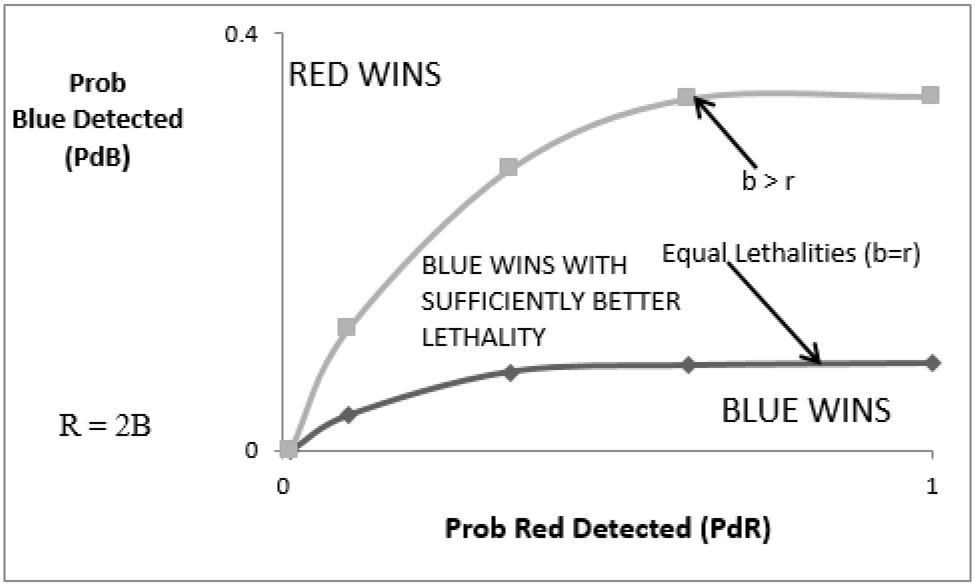


Figure 6. Stealth and Lethality in Overcoming Numerical Disadvantage

Program and Contract Schedule Uncertainty: Data Sources and Empirical Models

This section focuses on schedule changes and their associated cost and uncertainty during the Acquisition Process. Clearly, in light of the relationship between total cost growth and schedule and performance growth, it is necessary to explore in greater depth the information underlying total program acquisition cost. At the program level, this information is available in SARs; at a lower level, information is available in the Earned Value Management (EVM) system.

First SAR data are considered. The change in the acquisition cost associated with schedule constitutes one cost change among a specified group of program-level categories that experience cost changes. These changes underlie the total acquisition cost growth of Major Defense Acquisition Programs (MDAPs). The purpose of this analysis is to increase understanding of linkages among schedule categories within the acquisition process, and pave the way for a more complete understanding of the causes of schedule uncertainty.

It is helpful to begin with a brief discussion of total program cost growth, and then turn to the categories that identify cost changes—called "Program Variance" for a particular category. These categories are Program Schedule Variance, and Program Engineering Variance. Delving deeper, programs are typically underwritten using contracts. These contracts constitute the next lowest level for explaining the effects of schedule uncertainty. In fact, in each SAR report there are one or more contracts for which Earned Value Management data are reported on an annual basis. Annual data for these variances are contained in the Selected Acquisition Reports (SARs)

Several points should first be made about total program acquisition costs, which constitutes the top line for acquisition cost analysis and reporting. A significant portion of DoD Acquisition research has been devoted to the prospective breach of specified cost thresholds. Of particular importance are the thresholds mandated by Congress, which must be informed of significant breaches through Nunn-McCurdy Unit Cost Reporting.

These breaches would be monitored by the DAB, which establishes the Acquisition Defense Baseline (ADB). The ADB can change if a significant restructuring of the program occurs and there are new baseline cost estimates. A breach of the thresholds for both the Current Baseline Cost Estimate and the Original Baseline Cost Estimate would be of concern to both Congress and the acquisition community.

The Original Acquisition Program Baseline is determined at the time of Engineering and Manufacturing Development (EMD) approval (or Milestone B). These average costs are measured in relevant base year dollars (BY$): APUC equals ((Total Development + Procurement+ Construction) (in BY$)/Total Program Quantity), while PAUC equals (Total Procurement in BY$)/(Procurement Quantity).

It is important to understand that these cost growths are not directly associated with cost overruns that occur on contracts when the estimated final price rises above the current target price. Rather they result from a re-estimation of the effects of changes in the program parameters that are associated with the categories identified in the SAR reports, and, in turn, with the growth of total program acquisition costs.

Cost growth includes instances relating directly to schedule and indirectly to schedule via changes in the performance specifications. While specification changes can occur without changing the ABP, if there is a major restructuring of the program, the ABP may be changed by the DAB. This type of change could occur several times during the course of the acquisition process.

Under the purview of the Defense Acquisition Board (DAB), Performance and Schedule Breaches are also evaluated. These result from failures to meet specified Threshold Performance and Schedule values, where the Threshold values are typically lower than the Objective values.

Objective Schedule and Performance can be viewed as those levels that are “Best Value” to the government. These are the values that minimize full cost, which includes both accounting and implicit costs. Implicit costs might include costs of schedule slips.

Threshold Schedule and Performance, in contrast, refers the specified levels that are minimally acceptable to the government. With respect to Schedule, the delivery of different systems needs to be synchronized, and there is a cost associated with late delivery that affects this requirement. Also, if delivery is significantly late, a prior generation system may need to be retained in the inventory longer than anticipated at the same time as the effects of aging are incurred. There might also be an implicit cost from not achieving an Objective Performance level. In this case, minimum acceptable (Threshold) schedule and performance levels may be specified, with failure to achieve constituting a breach.

However, an examination of Selected Acquisition Reports indicates that frequently only Objective schedule and performance levels are specified. In this situation, a failure to meet Objective schedule milestones or the Objective levels of the relevant performance parameters would constitute a breach. In certain situations, however, a six-month delay in meeting a schedule milestone would constitute a breach.

Selected Acquisition Reports as Data Source

Program Variance information is reported in each annual SAR for selected categories. For each category, these variances are estimates of differences between the revised estimated final acquisition cost and the previous program development cost estimate. The development cost estimate is the total acquisition cost estimate developed at the time EMD is initiated.

Table 3 contains information specifying the different program variance categories identified in each SAR. Given these definitions, one can infer that the information reported for Program Engineering Variance and Program Schedule Variance reflects cost changes associated with contractually supported changes to the program. Both of these program variances can have a significant effect on the schedules that end up being achieved in the program. One might also expect Program Quantity Variance to have a significant effect on cost. However, there remains uncertainty as to how this program variance category should be analyzed, so in this preliminary analysis Program Quantity Variance will not be addressed. With respect to Program Economic Variance, the cost data are typically converted to constant dollar values by removing the effect of the price changes that Program Economic Variance embodies.

Calculations of cost breaches that fall under the Nunn-McCurdy requirements, would take account of the changes in cost associated with all of these SAR program variances, net of Program Economic Variance. The only difference would be that, for Nunn-McCurdy calculations the baseline cost would be that associated with the APB. For the SAR variances, the baseline would equal the program cost estimated at the time EMD is approved.

Table 3. SAR Program Variance Categories

Earned Value Management

To dig deeper, we turn to contractual data that underlies program data. Cost Variance (CV) is defined as follows:

$$CV = BCWP – ACWP, \tag{6}$$

where BCWP = Budget Cost of Work Performed and ACWP = Actual Cost of Work Performed. In turn, Schedule is defined as

$$SV = BCWP – BCWS, \tag{7}$$

where BCWS = Budgeted Cost of Work Scheduled.

These contract variances are depicted in Figure 7. In this case, the upper curve displays ACWP. Any time ACWP this is higher than BCWP, there is a cost overrun. The CV is therefore negative, which (counterintuitively) is typically an unattractive outcome. One measure of cost productivity would be ACWP/BCWP.

SV is somewhat more difficult to interpret. If BCWP is less than BCWS, then the scheduled work budgeted is greater than the work actually completed. That is, the work scheduled to be produced is less than the work actually produced, and the contract is behind schedule. A measure of schedule achievement productivity at this time is BCWP/BCWS.

Figure 7. Depicting Contract Variances

A question that impacts empirical work concerns the relationship between Program Variances and Contract Variances. As indicated, we restrict our focus to Program Engineering and Program Schedule Variance. The issue is how these relate to Contract Cost and Schedule Variance.

Worth noting is that each Program Variance is forward looking, and identifies estimated changes over the remainder of the program resulting (say) from specification changes. While Contract Variance is conceptually determined at a point in time, available data typically applies to some period of time such as a month, quarter, or year. It is for these time periods that cost and budget data are typically collected.

Empirical Analysis

Using the contract cost and budgetary data for a particular period does not seem to be an appropriate way to explain program variances that are associated with cost changes over the remainder of the program.

Also reported in the SARs and EVM documents are cumulative Program and Contract Variances. The cumulative program variances represent sums of projected cost changes since the date of the latest APB; the sums of contract variances represent aggregations of historical cost and budget information associated with work produced and work scheduled. These historical aggregations of **program and contract variance** overrun and underrun data are likely to be more effective in explaining the SAR program variance projections. Basically, if historical aggregations or budget and cost information indicate a tendency for outcomes to occur that are above budget and behind schedule, then, aggregations of SAR variance projections over time should be significantly related to the aggregation over time of contract variances. *Therefore, in the following empirical analysis, we use cumulative values of the program and contract variances.*

Two empirical results are displayed that result from the analysis of 31 SAR programs. In Table 4, Cumulative Program Engineering Variance is related to the two cumulative contract variances.

Table 4. Cumulative Engineering Variance Model

Explanatory Variable	Coefficient	t-Statistic	Significance
(Constant)	312.903	7.13	0.000
Cumulative Contract Schedule Variance	-6.194	-2.27	0.024
Cumulative Contract Cost Variance	-1.268	-1.91	0.057
Target Price Change	0.249	4.05	0.000
Dependent Variable: Cumulative Program Engineering Variance (2010 $M)			
R^2 = 0.136, N = 427			

After controlling for Target Price Change experienced by each contract, we find that both Cumulative Contract Schedule Variance are significantly negatively related to Cumulative Program Engineering Variance. The negative sign is plausible because increases in the contract variances reduce budget overruns and schedule slippage. Both of these reductions are shown to be associated with reductions in Cumulative Program Engineering Variance, that is, a decline in the cost growth of this program cost category.

The more negative coefficient of Cumulative Contract Schedule Variance suggests that contract schedule slippages are more likely to impact program performance specifications than contract outcomes in which actual cost is greater than budgetary cost.

In Table 5, after controlling for the relationship between EMD Achieved and Estimated EMD length, we find that both contract variances are negatively related to Program Schedule Variance. This is similar to the result shown in Table 4, and is consistent with expectations.

Once again, Contract Schedule Variance has a larger impact on Program Schedule Variance than Contract Cost Variance. Contract schedule slippages have a somewhat larger impact then being over contract cost expectations.

Table 5. Cumulative Program Engineering Variance Model

Explanatory Variable	Coefficient	t-Statistic	Significance
(Constant)	152.573	3.69	0.000
Cumulative Contract Schedule Variance	-5.303	-2.00	0.046
Cumulative Contract Cost Variance	-3.940	-7.17	0.000
Achieved - Estimated EMD Length	160.124	2.16	0.031
Dependent Variable: Cumulative Program Schedule Variance (2010 $M)			
R^2 = 0.174, N = 428			

Comments

One can probably reach an understanding of the linkages between program cost growth, SAR program-level cost variances, and contract-level data developed during the Earned Value Management process. However, the empirical relationships found between contract variances and both program engineering variance and program schedule variance require additional analysis. Typically cost variances are computed when there is no change in contract specifications. However, the SAR Engineering and Schedule Variances are dependent on those changes. Contract variances when they are negative may induce changes in performance specifications. However, this requires further analysis.

A useful next step is to focus in on the F-35 using a more detailed EVM data set. Explicitly included in the analysis would be schedule milestones and changes in the APB that have occurred. Hopefully, this will help answer some of the remaining open questions.

References

Abdel-Hamid, T. K., & Madnick, S. E. (1983). The dynamics of software project scheduling. *Communications of the ACM, 26*(5), 340–346. http://doi.org/10.1145/69586.358135

Airpower Australia. (2016, March). *Is the JSF really a fifth generation fighter?* Retrieved from http://www.ausairpower.net/jsf.html

Axe, D. (2011, December 13). Trillion-dollar jet has thirteen expensive new flaws. *WIRED*. Retrieved from https://www.wired.com/2011/12/joint-strike-fighter-13-flaws/

Axtell, L., & Irby, W. (2007, October 30–31). *APB, SAR, and Nunn-McCurdy Status*. DAMIR Conference Briefing.

Balaji, J., & James, R. B. (2005). *Project dynamics with applications to change management and earned value tracking.* Presented at the System Dynamics Society International Conference.

Bogdan, C. (2012, September 17). *Joint strike fighter requirements*. Remarks to Air Force Association Air & Space Conference, National Harbor, MD. Retrieved from http://www.af.mil/shared/media/document/AFD-120919-046.pdf

Bolten, Joseph, et al. (2008). *Sources of weapons system growth: Analysis of 35 major defense acquisition programs*. Santa Monica, CA: RAND.

Bracken, J., Kress, M., & Rosenthal, R. (Eds.). (1995). *Warfare modeling*. Hoboken, NJ: Wiley.

CJCS. (2015). *Joint capabilities integration and development system (JCIDS)* (CJCS Instruction 3170.01I). Washington, DC: Author. Retrieved from https://dap.dau.mil/policy/Documents/2015/CJCSI_3170_01I.pdf

Cooper, K. G. (1980). Naval ship production: A claim settled and a framework built. *Interfaces, 10*(6), 20–36. http://doi.org/10.1287/inte.10.6.20

Cooper, K. G. (1993a). The rework cycle: Benchmarks for the project manager. *Project Management Journal*, XXIV(1), 17–22.

Cooper, K. G. (1993b). The rework cycle: How it really works … and reworks. *PM Network,* 25–28.

Cooper, K. G. (1993c). *The rework cycle: Why projects are mismanaged*. Project Management Institute.

Cooper, K. G., & Mullen, T. W. (1993). Swords and plowshares: The rework cycles of defense and commercial software development projects. *American Programmer*.

Cooper, K. G., Lyneis, J. M., & Bryant, B. J. (2002). L*earning to learn, from past to future, 20*(3), 213–219. http://doi.org/10.1016/S0263-7863(01)00071-0

Coyle, R. G. (1996). *System dynamics modelling.* Boca Raton, FL: CRC Press.

Defense Science Board (DSB). (1994). Task Force on Joint Advanced Strike Technology (JAST) program. Retrieved from http://www.dtic.mil/docs/citations/ADA292094

DoD. (n.d.-a). JSF history, Introduction. Retrieved from http://www.jsf.mil/history/index.htm

DoD. (n.d.-b). JSF history, JAST. Retrieved from http://www.jsf.mil/history/his_jast.htm

DoD. (n.d-c). JSF history, Pre-JAST. Retrieved from http://www.jsf.mil/history/his_prejast.htm

Director, Operational Testing & Evaluation (DOT&E). (2016). *DOT&E FY2016 annual report*: F-*35 Joint Strike Fighter*. Retrieved from http://www.dote.osd.mil/pub/reports/FY2016/pdf/dod/2016f35jsf.pdf

Forrester, J. W. (1987). Lessons from system dynamics modeling. *System Dynamics Review, 3*(2), 136–149. http://doi.org/10.1002/sdr.4260030205

Forrester, J. W. (1995). The beginning of system dynamics. *McKinsey Quarterly*.

Franck, R., Hildebrandt, G., & Udis, B. (2016). Toward realistic cost schedule estimates. In *Proceedings of the 13th Annual Acquisition Research Symposium* (Vol. I, pp. 95–116). Monterey, CA: Naval Postgraduate School. Retrieved from http://www.acquisitionresearch.net/files/FY2016/SYM-AM-16-029.pdf

Fraoili, S. A. (2016, Summer). Intelligence support for the F-35A Lightning II. *Air and Space Power Journal, 30*(2), 106–10 Retrieved from http://www.au.af.mil/au/afri/aspj/digital/pdf/issues/2016/ASPJ-Summer-2016.pdf

GAO. (2015). *Acquisition reform: DOD should streamline its decision-making process for weapon systems to reduce inefficiencies* (GAO-15-192). Retrieved from http://www.gao.gov/products/GAO-15-192,February

GAO. (2016a). *F-35 joint strike fighter: Continued oversight needed as program plans to begin development of new capabilities* (GAO-16-390). Retrieved from http://www.gao.gov/assets/680/676584.pdf

GAO. (2016b). *Weapon system requirements: Detailed systems engineering prior to product development positions programs for success* (GAO-17-77). Retrieved from http://www.gao.gov/products/GAO-17-77

Godlewski, E., Lee, G., & Cooper, K. (2012). System dynamics transforms Fluor project and change management. *Interfaces, 42*(1), 17–32. http://doi.org/10.1287/inte.1110.0595

Haupt, E. (2007). *BAS105 basic earned value management* [Briefing).

Herroelen, W. (2005). Project scheduling—Theory and practice. *Production and Operations Management.*

Hildebrandt, G. G., & Sze, M. (1986, July). *Accounting for the cost of tactical aircraft* (N-2420 -PA&E). Santa Monica, CA: RAND. Retrieved from http://www.rand.org/content/dam/rand/pubs/notes/2007/N2420.pdf

Hough, P. G. (1996). *Pitfalls in calculating cost growth for selected acquisition programs*. Santa Monica, CA: RAND

Hughes, R. T. (1996). Expert judgement as an estimating method. *Information and Software Technology, 38*(2), 67–75. http://doi10.1016/0950-5849(95)01045-9

Kerzner, H. R. (2013). *Project management: A systems approach to planning, scheduling, and controlling* (11th ed). Hoboken, NJ: Wiley.

Laird, R. (2012, September 11). The fifth generation experience updated: The F-35 is a situational awareness machine. SLDInfo.com. Retrieved from http://www.sldinfo.com/the-fifth-generation-experience-updated-the-f-35-is-a-situational-awareness-machine/

Lockie, A. (2017). F-35 pilot: Here's what people don't understand about dogfighting, and how the F-35 excels at it. *Business Insider.* Retrieved from http://www.businessinsider.com/f35-pilot-f-35-can-excel-dogfighting-2017-1

Lorell, M. A., et al. (2015). *Extreme cost growth: Themes from six U.S. Air Force major defense acquisition programs.* Santa Monica, CA: RAND.

McNicol, D. L. (2005). *Cost growth in major weapon procurement programs*. Washington, DC: Institute for Defense Analysis.

Miller, T. (2016, December 6). The moment pilots first realized the F-35 was something extraordinary. SLDInfo.com. Retrieved from http://www.sldinfo.com/the-moment-pilots-first-realized-the-f-35-was-something-extraordinary/

Moder, J. J., Phillips, C. R., & Davis, E. W. (1983). *Project management with CPM, PERT and precedence diagramming* (3rd ed.). New York, NY: Van Nostrand Reinhold.

Reichelt, K., & Lyneis, J. (1999). The dynamics of project performance: Benchmarking the drivers of cost and schedule overrun. *European Management Journal, 17*(2), 135–150. http://doi.org/10.1016/S0263-2373(98)00073-5

Rodrigues, A., & Bowers, J. (1996). System dynamics in project management: A comparative analysis with traditional methods. *System Dynamics Review, 12*(2), 121–139. http://doi.org/10.1002/(sici)1099-1727(199622)12:2<121::aid-sdr99>3.3.co;2-o

Seligman, L. (2017, March 28). How Tiltrotor Tech is changing the way the U.S. Marine Corps fights. *Aviation Week*. Retrieved from http://aviationweek.com/defense/how-tiltrotor-tech-changing-way-us-marine-corps-fights

Steidle, C. (1996, December 31). *JSF Selected Acquisition Report*. Washington, DC: DoD.

Steidle, C. E. (1997, January–March). The joint strike fighter program. *Johns Hopkins APL Technical Digest (18)*1, 6–20. Retrieved from http://www.dtic.mil/docs/citations/ADA518821-223KB-2006-06-02PDF

Sterman, J. (2000). *Business dynamics: Systems thinking and modeling for a complex world with CD-ROM*. New York, NY: McGraw-Hill Education.

Sweetman, B. (2012, November 26). Generation gap. *Aviation Week, 22–23* Retrieved from http://archive.aviationweek.com/issue/20121126#!&pid=22

Sydney J. (2017, February 8). Navy, Marine F-18s in "death spiral" as readiness plummets, *Breaking Defense*. Retrieved from http://breakingdefense.com/2017/02/navy-marine-f-18s-in-death-spiral-as-readiness-plummets/

Taylor, J. G. (1983). *Lanchester models of warfare* (two volumes). Military Applications Section, Operations Research Society of America.

Tirpak, J. A. (2017a, February 24). Network news. *Air Force Magazine Daily Report*. Retrieved from http://www.airforcemag.com/DRArchive/Pages/2017/February%202017/February%2024%202017/Network-News.aspx

Tirpak, J. A. (2017b, March 23). The big dig. *Air Force Magazine Daily Report*. Retrieved from http://www.airforcemag.com/DRArchive/Pages/2017/March%202017/March%2023%202017/The-Big-Dig.aspx

Tirpak, J. A. (2017c, April–May). Saving air superiority. *Air Force Magazine, 100*(4), 24. Retrieved from http://www.airforcemag.com/MagazineArchive/Pages/2017/April%202017/Saving-Air-Superiority.aspx

Weisgerber, M. (2016). US Air Force officers give new details for F-35 in war with China. In B. Peniston (Ed.), *What's next for military aviation*. Retrieved from http://www.defenseone.com/assets/military-aviation-ebook/portal/

Williams, T. (2002). *Modelling complex projects*. New York, NY: John Wiley & Sons.

Younossi, O., et al. (2007). *Is cost growth increasing: A quantitative assessment of completed and ongoing programs*. Santa Monica, CA: RAND.

THIS PAGE INTENTIONALLY LEFT BLANK

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CA 93943

www.acquisitionresearch.net