Acquisition Research: creating synergy for informed change; May 13-14, 2015

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/47664
PROCEEDINGS
OF THE
TWELFTH ANNUAL ACQUISITION RESEARCH SYMPOSIUM

THURSDAY SESSIONS
VOLUME II

Acquisition Research:
Creating Synergy for Informed Change

May 13–14, 2015

Published April 30, 2015

Disclaimer: The views represented in this report are those of the author and do not reflect the official policy position of the Navy, the Department of Defense, or the federal government.
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

Keynote: The Honorable Sean Stackley, Assistant Secretary of the Navy, Research, Development, and Acquisition ................................................................. iii

Plenary Panel: Reducing Lifecycle Sustainment Costs ........................................ 1


- Achieving Better Buying Power Through Acquisition of Open Architecture Software Systems for Web-Based and Mobile Devices ........................................ 4
- DoD Software-Intensive Systems Development: A Hit and Miss Process ........ 22
- Investing in Software Sustainment ................................................................... 41

Panel 13. DoD’s Contracting Portfolio: Analysis and Innovations ..................... 49

- Contractor Past Performance Information: An Analysis of Assessment Narratives and Objective Ratings ................................................................. 69


- Analytical Tools for Affordability Analysis ....................................................... 83
- Government Contractors and Sticky SGA Costs ................................................ 95


- Addressing the Barriers to Agile Development in the Department of Defense: Program Structure, Requirements, and Contracting ............................. 110
- Finding the “RITE” Acquisition Environment for Navy C2 Software ............. 127
- Use of Automated Testing to Facilitate Affordable Design of Military Systems .. 139

Panel 16. Managing Obligations Rates, Expenditures, and Internal Controls in the DoD ............................................................................................................. 152

- OSD’s Obligation & Expenditure Rate Goals: An Examination of the Factors Contributing to the Interference ......................................................... 153
- Contracting Processes, Internal Controls, and Procurement Fraud: A Knowledge Assessment ......................................................................................... 169
- Department of Defense Management of Unobligated Funds for Acquisition Programs ..................................................................................................... 186

Panel 17. Applying Portfolio Techniques to Acquisition Management .............. 200

- An Analytic Model for DoD Divestments .......................................................... 228
Portfolio Acquisition—How the DoD Can Leverage the Commercial Product Line Model ................................................................. 249

Panel 18. Promoting Secure and Timely Development and Sustainment in Weapons System Procurement ................................................................. 267
  Towards Rapid Recertification Using Formal Analysis ................................................................. 269

Panel 19. Factors Influencing the Selection of Contracting Strategies ........ 289
  Antecedents and Consequences of Federal Bid Protests ................................................................. 291
  Avoiding Terminations, Single Offer Competition, and Costly Change Orders with Fixed-Price Contracts ................................................................. 307
  Price Analysis on Commercial Item Purchases Within the Department of the Navy ................................................................. 326

Panel 20. Understanding the Relevance of Big Data and Predictive Analytics in Business and the Defense Industrial Base ......................................................... 343
  Lexical Link Analysis Application: Improving Web Service to Acquisition Visibility Portal Phase III ................................................................. 345

Additional Papers ....................................................................................................................................... 364
  Evaluating Intellectual Property and Data Rights in Competition Source Selections—Leveraging the “Assertions Process” to a New Level to Foster Open Systems Architecture ................................................................. 365
  Orchestrating the Development of a Complex System of Systems: Systems Engineering Tools and Methodologies ................................................................. 386
Keynote: The Honorable Sean Stackley, Assistant Secretary of the Navy, Research, Development, and Acquisition

The Honorable Sean J. Stackley—assumed the duties of assistant secretary of the Navy for Research, Development, and Acquisition (ASN [RDA]) following his confirmation by the Senate in July 2008. As the Navy’s acquisition executive, Stackley is responsible for the research, development, and acquisition of Navy and Marine Corps platforms and warfare systems, which includes oversight of more than 100,000 people and an annual budget in excess of $50 billion.

Prior to his appointment to ASN (RDA), Stackley served as a professional staff member of the Senate Armed Services Committee. During his tenure with the Committee, he was responsible for overseeing Navy and Marine Corps programs, U.S. Transportation Command matters, and related policy for the Seapower Subcommittee. He also advised on Navy and Marine Corps operations and maintenance, science and technology, and acquisition policy.

Stackley began his career as a Navy surface warfare officer, serving in engineering and combat systems assignments aboard USS John Young (DD 973). Upon completing his warfare qualifications, he was designated as an engineering duty officer and served in a series of industrial, fleet, program office, and headquarters assignments in ship design and construction, maintenance, logistics, and acquisition policy.

From 2001 to 2005, Stackley served as the Navy’s LPD 17 program manager, with responsibility for all aspects of procurement for this major ship program. Having served earlier in his career as production officer for the USS Arleigh Burke (DDG 51) and project naval architect overseeing structural design for the Canadian Patrol Frigate, HMCS Halifax (FFH 330), he had the unique experience of having performed a principal role in the design, construction, test, and delivery of three first-of-class warships.

Stackley was commissioned and graduated with distinction from the United States Naval Academy in 1979, with a Bachelor of Science in mechanical engineering. He holds the degrees of Ocean Engineer and Master of Science, mechanical engineering, from the Massachusetts Institute of Technology. Stackley earned certification as a professional engineer, Commonwealth of Virginia, in 1994.
Plenary Panel: Reducing Lifecycle Sustainment Costs

Thursday, May 14, 2015

| 9:30 a.m. – 11:00 a.m. | **Chair:** The Honorable David J. Berteau, Assistant Secretary of Defense Logistics and Materiel Readiness

**Panelists:**
- Scott DiLisio, Director, Strategic Mobility/Combat Logistics Division, Chief of Naval Operations
- Richard Burke, Director, Operations Analysis & Procurement Planning Division, Office of the Secretary of Defense, Cost Assessment and Program Evaluation (CAPE)
- Daniel A. Fri, Deputy Assistant Secretary of the Air Force for Logistics and Product Support, Office of the Assistant Secretary of the Air Force for Acquisition

The Honorable David J. Berteau—is the assistant secretary of defense for logistics and materiel readiness, responsible for ensuring world class military logistics support to the men and women of the United States Armed Forces. He manages logistics policy and processes to provide superior, cost-effective, joint logistics support to the warfighter. He oversees the management of $170 billion in Department of Defense logistics operations.

Prior to this role, Berteau served as senior vice president and director of the National Security Program on Industry and Resources at the Center for Strategic and International Studies (CSIS) in Washington, DC. His research and analysis covered national security plans, policies, programs, budgets, and resources; defense management, contracting, logistics, and acquisition; and national security economics and industrial base issues. Berteau has been an adjunct professor at Georgetown University and at the Lyndon B. Johnson School of Public Affairs, a director of the Procurement Round Table, and an associate at the Robert S. Strauss Center at the University of Texas. He remains a fellow of the National Academy of Public Administration.

Before he joined CSIS full time in 2008, he served as a CSIS non-resident senior associate for seven years. In addition, he was director of national defense and homeland security for Clark & Weinstock, director of Syracuse University's National Security Studies Program and a professor of practice at the Maxwell School of Citizenship and Public Affairs, and senior vice president at Science Applications International Corporation (SAIC). Previously, he served 12 years at senior levels in the U.S. Defense Department under four defense secretaries, including four years as principal deputy assistant secretary of defense for production and logistics.

Berteau graduated with a BA from Tulane University in 1971 and received his master's degree in 1981 from the LBJ School of Public Affairs at the University of Texas.

Thursday, May 14, 2015

11:15 a.m. – 12:45 p.m.

Chair: Captain Kurt Rothenhaus, USN, Commanding Officer, SPAWAR Systems Center Pacific

Achieving Better Buying Power Through Acquisition of Open Architecture Software Systems for Web-Based and Mobile Devices
   Walt Scacchi, University of California–Irvine
   Thomas Alspaugh, University of California–Irvine

DoD Software-Intensive Systems Development: A Hit and Miss Process
   Brad Naegle, NPS

Investing in Software Sustainment
   Robert Ferguson, Software Engineering Institute

Captain Kurt Rothenhaus, USN—is a native of New York City and 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; he 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 15, and Combat Systems/C5I officer on USS Harry S. Truman (CVN-75). Additionally, he served in Baghdad, Iraq, developing counter-insurgency and reconstruction systems to the Army Corps of Engineers in 2006.

His acquisition assignments include serving as project manager for Space and Naval Warfare System Center Pacific, assistant program manager for Maritime Domain Awareness in PMW-120, Pacific Fleet CVN modernization manager in PMW-750, Future Command and Control assistant program manager in PMW-150, assistant program manager for the Consolidated Afloat Network Enterprise System network program in PMW-160, and deputy program manager for communications and GPS in PMW/A-170.

He has been recognized with the A. Bryan Laswell National Defense Industrial Association Award in 2007 for technology innovation and a 2008 Navy & Marine Corps Leadership Award while serving aboard CVN-75.

His personal awards include the Meritorious Service Medal, Joint and Navy Commendation Medal, Navy Achievement medals, and various service and campaign awards.
Achiving Better Buying Power Through Acquisition of Open Architecture Software Systems for Web-Based and Mobile Devices

Walt Scacchi—is senior research scientist and research faculty member at the Institute for Software Research, University of California, Irvine. He received a PhD in information and computer science from UC Irvine in 1981. From 1981–1998, he was on the faculty at the University of Southern California. In 1999, he joined the Institute for Software Research at UC Irvine. He has published more than 150 research papers, and has directed more than 65 externally funded research projects. In 2011, he served as co-chair for the 33rd International Conference on Software Engineering—Practice Track, and in 2012, he served as general co-chair of the 8th IFIP International Conference on Open Source Systems. [wscacchi@ics.uci.edu]

Thomas Alspaugh—is a project scientist at the Institute for Software Research, University of California, Irvine. His research interests are in software engineering, requirements, and licensing. Before completing his PhD, he worked as a software developer, team lead, and manager in industry, and as a computer scientist at the Naval Research Laboratory on the Software Cost Reduction, or A-7, project. [alspaugh@ics.uci.edu]

Abstract

Many people within large enterprises rely on up to four Web-based or mobile devices for their daily work routines—personal computer, tablet, and personal and work-specific smartphones. Our research is directed at identifying, tracking, and analyzing software component costs and cost reduction opportunities within the acquisition life cycle of open architecture (OA) systems for such Web-based and mobile devices. These systems are subject to different intellectual property license and cybersecurity requirements. Our research goal is to create a new approach to address challenges in the acquisition of software systems for Web-based or mobile devices used within academic, business, or government enterprises. Acquisition personnel in such enterprises will increasingly be called on to review and approve choices between functionally similar open source software (OSS) components, and commercially priced closed source software (CSS) components, to be used in the design, implementation, deployment, and evolution of secure OA systems. We seek to make this a simpler, more transparent, and more tractable process. Finally, this acquisition research supports and advances a public purpose by investigating acquisition challenges arising from the adoption and deployment of secure OA software systems for Web-based or mobile devices.

Overview

The Department of Defense (DoD), other government agencies, and most large-scale business enterprises continually seek new ways to improve the functional capabilities of their software-intensive systems with lower acquisition costs. The acquisition of open architecture (OA) systems that can adapt and evolve through replacement of functionally similar software components is an innovation that can lead to lower cost systems with more powerful functional capabilities. OA system acquisition, development, and deployment are

---

1 This report was supported by the Acquisition Research Program at the Naval Postgraduate School, Monterey, CA. No endorsement, review, or approval implied. This paper reflects the views and opinions of the authors, and not necessarily the views or positions of any other persons, group, enterprise, or government agency.
thus seen as an approach to realizing Better Buying Power (BPP) goals for lowering system costs, achieving technical excellence, enabling innovation, and advancing the acquisition workforce.

Our research identifies and analyzes how new software component technologies like apps and widgets for Web-based and/or mobile devices, along with their intellectual property (IP) license and cybersecurity requirements interact to drive down (or drive up) total system costs across the system acquisition life cycle. The availability of such new scientific knowledge and technological practices can give rise to more effective expenditures of public funds and improve the effectiveness of future software-intensive systems used in government and industry. Thus, a goal of this presentation is to explore new ways and means for achieving cost-sensitive acquisition of OA software systems, as well as identifying factors that can further decrease or increase the costs of such systems at this time.

We begin by briefly reviewing to identify a set of recent trends in the development of OA software systems that intend to develop more capable OA systems. These trends include the transition to adoption of small-form factor software components as distinct applications (standalone and plug-in “apps”) and widgets that exploit modern Web capabilities. We then turn to examine some key goals of the BBP 3.0 initiative (Kendall, 2014) that direct attention to adoption of OA system development practices that affect acquisition practices. Next, we identify a new set of emerging challenges to achieving BBP through OA software systems. We then identify three new practices to realize the cost-effective acquisition of OA software systems.

Recent Trends Affecting Better Buying Power Through OA Systems

We find there are four broad trends that mediate the cost-effectiveness and buying power of emerging OA system acquisition efforts. These include (a) the move towards shared, multi-party acquisition and agile development of new OA systems across compatible software ecosystems; (b) exploitation of new software component technologies compatible with Web and mobile devices; (c) growing diversity of cybersecurity challenges to address during system development; and (d) new software development business models for app/widget development and deployment. Each is examined in turn.

A. Multi-Party Acquisition and Development System Ecosystems

Many in the defense community seek to embrace the acquisition and development of agile command and control (C2) and related enterprise systems (Agre et al., 2014; George, Bowers, et al., 2014; George et al., 2013; Guertin & Womble, 2012; Reed et al., 2012; Scacchi & Alspaugh, 2012b, 2013c, 2014a). Such systems are envisioned to arise from the assembly and integration of system elements (application components, widgets, content servers, networking elements, etc.) within a software ecosystem of multiple producers, integrators, and consumers who may supply or share the results of their efforts. The assembly and integration of system elements produces “assembled capabilities for C2 systems” (AC-C2). Our purpose is to identify how our approach to the design of secure OA systems can be aligned with this emerging vision for agile C2 system development and adaptive deployment. We also focus on design of OA system capability involving office productivity and social media components (Agre et al., 2014) that increasingly may be configured within a secure AC-C2 (Scacchi & Alspaugh, 2011, 2012b, 2013b).

The design and development of agile C2 systems follows from two sets of principals: one set addressing guidelines/tenets for multi-party engineering (MPE) of C2 system components; the other set addressing attributes of agile and adaptive ecosystems (AAE) for producing AC-C2s or C2 system elements (Reed et al., 2012; Reed et al., 2014; Scacchi &
Alspaugh, 2014a, 2014b, 2014c). To help understand what we mean by a *software ecosystem*, we use Figure 1 to represent where different parties are located across a generic software supply networks or multi-party relationships that emerge to enable the software producers to develop and release products that are assembled and integrated by system integrators for delivery to end-user organizations, via online storefronts (George, Bowers, et al., 2014; George, Galdorisi, et al., 2014; George et al., 2013).

As noted, OA system components can include software applications (apps) and widgets. Widgets are lightweight, single-purpose web-enabled applications that users can configure to their specific needs (Agre et al., 2014; Gizzi, 2011; George et al., 2013; Scacchi & Alspaugh, 2013b). Widgets can provide summary information or a limited view into a larger application that can be used alongside related widgets to provide an integrated view, as required by users.

The lower part of Figure 1 also identifies where elements of shared agreements like IP licenses or cybersecurity requirements enter into the ecosystem, and how the assembly of components into a configured system or subsystem architecture by system integrators effectively (and perhaps unintentionally) determines which IP license or cybersecurity obligations and rights get propagated to consumer or end-user organizations. Agreement terms and conditions acceptable to consumer/end-user organizations flow back to the integrators. This helps reveal where and how shared agreements will mix, match, mashup, or encounter semantic mismatches at the system architecture level, which is one reason why we use (and advocate) explicit OA system models.

Overall, a move towards MPE and AAE substantiates a path towards decentralized OA system development, integration, and deployment (DoD, 2012; Gizzi, 2011).
This decentralization will engender acquisition and development of heterogeneously-licensed systems (HLS), whereby different software components (apps, widgets) will be subject to different IP licenses (Alspaugh et al., 2012; Alspaugh et al., 2010), as well as to different cybersecurity requirements (Defense Acquisition Guidebook, 2014; Scacchi & Alspaugh, 2012b, 2013a, 2013b, 2013c). This implies that such components, their IP licenses, and cybersecurity requirements will be subject to ongoing evolution across a diversity of methods, shown in Figure 2 (Scacchi & Alspaugh, 2012a, 2013b). These will create a new generation of challenges for the acquisition workforce, in terms of training, new work and contract management practices, and need for automated assistance to track and manage oversight of policy compliance (e.g., for alignment with BPP and cybersecurity assessment). Without automated assistance, it appears that the acquisition workforce will be
overwhelmed with technical details that interact with acquisition, development, and/or system integration contracts and software component IP licenses and cybersecurity requirements. Otherwise, these conditions suggest that acquisition management practices can complicate acquisition (George, Bowers, et al., 2014), and thus potentially mitigate the benefits of BBP that can arise from MPE and AAE for C2 systems.

Figure 2. The Kinds of Common Evolutionary Changes That Arise During OA Software Component Development, Deployment, and Sustained Usage

Moving Towards Shared Development of Apps and Widgets as OA System Components

Future OA systems for agile C2 may be configured by system integrators, end-user organizations, or warfighters in the field. This would be accomplished through access to online repositories of software apps or user-interface widgets. The Ozone Widget Framework (OWF) is a government open source software (GOSS) effort that is central to such agile OA system development. The OZONE family of products includes the OWF and the OZONE Marketplace, the marketplace being an online repository whose operation is similar in kind to the online app stores by Apple and Google (Scacchi & Alspaugh, 2013b). These products are built to fit the needs of human centered fusion activities in network-centric warfare environments. The OZONE family of products is designed as a presentation layer toolkit that can be rapidly deployed in a variety of mission contexts ranging from strategic planning to enable the creation of a real-time common operational picture and situation awareness applications. Figure 3 displays examples of OWF-based widgets operating in a Web browser, while Figure 4 shows OWF widgets deployed for use on a mobile device.

Growing Diversity of Challenges in Cybersecurity

New types of software components like apps and widgets must be developed, deployed, and sustained in ways compatible with existing cybersecurity requirements. They must also be later adapted to accommodate emerging cybersecurity requirements that are not yet apparent. For example, there is growing interest in accommodating not just mobility, but also “Bring Your Own Device” (BYOD) capabilities.
BYOD suggests that end-users and warfighters are bringing their own mobile devices with them into the field to support their missions. However, BYOD clearly exacerbates the technical challenges of cybersecurity assurance, often in ways that cannot be readily anticipated, as when independently developed components co-evolve in conflict to one another (Weir, 2014). Nonetheless, acquisition policy necessitates that cybersecurity vulnerability and exposures be addressed (Defense Acquisition Guidebook, 2015). But at present, it is unclear what new kinds of requirements these new OA system components
bring to the acquisition workforce. For example, a move to adopt mobile apps and/or mobile widgets means these OA system components must pass through an application security process for “vetting” these components.

Vetting entails establishing what cybersecurity requirements are to be verified, how they are to be validated, as well as where, when, and by whom these activities should be performed. One approach is to assume the vetting can be performed by a centralized authority, such as by the operator of the Ozone Marketplace. But it is not clear that there will ever be only one such authority.

Instead, if we foresee multiple marketplaces, which are already appearing both in GOSS and industrial online settings, then the acquisition workforce will be challenged in how best to determine which cybersecurity requirements must be addressed, validated, and compliance certified, as well as by whom and how often. Consider the example, seen in Figure 5, of a widget for “emergency response incident command system,” developed for the Department of Homeland Security (Rockwell, 2015). How do its components (possibly GOSS) compare or interoperate with widgets/AC-C2 from DoD agencies or program offices concerned with C2 system interoperability or AC-C2?

Figure 5. AC-C2 Style Widget From the Next-Generation Incident Response System for DHS

A move to widgets also presents new kinds of cybersecurity challenges when two or more widgets are configured together with one or more apps to create a mashup that provides an agile system capability. This situation refers to the technical challenges of inter-widget communication. Such component–component communication can be technically realized in different ways, such as via ad hoc, “open standards,” or publish–subscribe messaging interfaces, as well as whether point-to-point or as configured through a dynamic processing mashup (Chudnovsky et al., 2013; Endres-Niggemeyer, 2013a). While OA systems may rely on “open standards”–style widget interfaces and communications patterns
may be used, widget communication/interface standards/interfaces are still very new technologies and techniques. Thus, it is unclear which will survive and be widely adopted (Endres-Niggemeyer, 2013b).

Similarly, knowledge about the proper usage of widget components is unclear, and thus is not yet ready for compliance assessment within current acquisition practices. The technical challenge is further complicated when apps/widgets are acquired from different online marketplaces. Different marketplaces may rely on different schemes for specification and interchange of shared data semantics between autonomously developed components. This in turn hinges on the expertise of OA system integrators, end-users, or warfighters to recognize how, where, and when the semantics of technical data interchange arise and to what consequences via component–component API alignments (to avoid mismatches), data type representations, data formats (e.g., "CSV" vs. .xls vs. XML), data naming conventions (for resource discovery vs. data modeling ontology), data range value limits, exceptional values, data-flow control signals, and so forth. These are still new technical problems that are yet to be readily resolved or to have development/usage guides.

**New Business Models for OA Software Component Development and Use**

New business models imply differentiated IP licenses and contracting practices. Given our discussion up to this point, along with reference to our recent acquisition research studies (Alspaugh et al., 2012; Scacchi & Alspaugh, 2011, 2012b, 2013b), this means different obligations and rights will be transferred from component producers to system integrators and end-user organizations. Some licenses are “buy and pay now,” while others are “free now, pay later, based on usage,” others are “many organizations (e.g., PEOs) will share purchase costs,” and so forth.

Acquisitions of new kinds of OA system components allow for new business models. These include new models for software component producers, system integrators, and end-user organizations. For example, new software and OA system development business models for software app/widget development and deployment include (in no particular order) the following: (1) franchising, (2) enterprise licensing, (3) metered usage, (4) advertising supported, (5) subscription, (6) free component, (7) paid service/support fees, (8) federation reciprocity for shared development, (9) collaborative buying, (10) donation, (11) sponsored development, (12) free/open source software (e.g., Government OSS [GOSS]), and others (D. Hanf, personal communication, July 2013). Further, this list is not exhaustive; instead, it is only representative.

In contrast, for end-user organizations involved in agile development of OA system components, or an integrated system capability, there is a need to develop and codify their own business models regarding OA software component development or system integration. These business models are constituted through “shared agreements” that allow for sharing the cost of component or integrated capability development and cybersecurity assurance vetting across multiple parties (e.g., multiple program offices). However, these shared agreements are also a core part of emerging MPE/AAE development practices. These agreements must convey how OA component development or system integration costs and security assurance will be shared, as well as how they will be sustained in the presence of interacting software component development, deployment, and evolution processes and practices (Scacchi & Alspaugh, 2013a). Shared agreements denote the obligations the participating organizations are willing to accept, in order to realize the provided rights they need. So shared agreements can be expressed and assessed in the same manner, and with the same analysis tools and techniques, as IP licenses and cybersecurity requirements (Scacchi & Alspaugh, 2013b, 2013c).
Software acquisition costs easily become difficult to predict/manage given the diversity of business models, IP licenses, and implied software component cybersecurity assessment. Development/usage cost sharing agreements can further complicate determination of development cost, costs shares across organizations, and system costs over time as business models, component licenses, and cybersecurity assessment requirements evolve (Scacchi & Alspaugh, 2012a, 2013a).

What kind of expertise do we expect the acquisition workforce to need in order to make adoption of “component-based system capabilities” (including for mobile devices) agile, adaptive, and practical across different commercial/governmental software marketplaces/ecosystems? What kinds of acquisition guidance is needed for articulating and streamlining Shared Agreements between multiple organizations participating in shared OA component development and cybersecurity assurance? What kinds of acquisition management practices and analysis tools are needed for the acquisition workforce to ensure cost savings and BBP in such settings? Addressing these questions is beyond the scope of this paper, but these questions require follow-on acquisition research to resolve and answer.

Better Buying Power 3.0 Goals

Better Buying Power (http://bbp.dau.mil/) is part of the DoD’s initiative that sees continuous improvement as the best approach to improving the performance of the defense acquisition enterprise. BBP 3.0 (Kendall, 2014) identifies eight areas of focus that group a larger set of itemized initiatives that offer the potential to restore affordability and realize technical excellence in defense procurement and improve defense industry productivity. One of the eight areas focuses on promoting or increasing competition, and this area includes an initiative to utilize modular open system architectures to stimulate innovation (Kendall, 2014). Technical innovations are constrained by two categories of Intellectual Property (IP) rights available to the government: (a) technical data (TD; e.g., product design data, computer databases, computer software documentation) and (b) computer software (CS; e.g., source code, executable code, design details, processes, and related materials). These rights are realized through IP licenses provided by system product or service providers (e.g., software producers) to the government customer, so long as the customer fulfills the obligations stipulated in the license agreement (e.g., to indicate how many software users are authorized to use the licensed product or service according to a fee paid).

As already noted, our acquisition research has focused on issues addressing OA systems and IP licenses since 2008 (Scacchi & Alspaugh, 2008), as well as forward to the acquisition of secure OA systems for command and control (C2) and enterprise information systems (Scacchi & Alspaugh, 2011, 2012b, 2013b), where security requirements can be expressed in a manner similar to IP obligations and rights. Therefore, here we turn to identify how a sample of different goals of BBP 3.0 initiatives interact or relate to the trends and challenges examined so far in this paper. The BBP goals are highlighted, and then followed by a brief examination.

- Promote effective competition—One central purpose for acquiring OA systems is to increase the likelihood of creating and maintaining competitive environments among system producers who can provide software components that can be replaced by similar offerings by other component producers. We demonstrate how this can work when system architectures are explicitly modeled, and their software components and interconnections are similarly specified in an open manner (Alspaugh et al., 2012; Scacchi & Alspaugh, 2012a). Such openness also supports improved technology search.
Use Modular Open Systems Architecture to stimulate innovation—Open system architectures that can accommodate common components from alternative producers requires that the components utilize standardized interfaces, whether in the form of open Application Program Interfaces (APIs), standard data exchange protocols, and standard data representations, formats, and meta-data, as well as utilization of open source software (OSS) components (Scacchi & Alspaugh, 2008). But as also noted earlier, app and widget components at present have a plethora of standardized interfaces, and it is unclear which will survive, be sustained, be widely adopted (inside/outside of the DoD), and be evolved (Endres-Niggemeyer, 2013b).

Increase small business participation and opportunities—One way to increase competition in the realm of OA systems is to identify where smaller scale software applications (apps) or widgets can be utilized, which might be produced by innovative small businesses or startup ventures which dominate much of the online markets for Web-based or mobile device apps/widgets. Small businesses may further be advantaged by their utilization of shared OSS infrastructure components, platforms, or remote services, since large commercial contractors may not see sufficient profit margins to develop proprietary alternatives. So OA systems that accommodate OSS components that can integrate custom apps/widgets into innovative system capabilities (AC-C2), may then realize new opportunities for DoD customers. Other small business opportunities may similarly arise for such ventures that focus on emerging cybersecurity assessment or tool development services.

Improve our leaders’ ability to understand and mitigate technical risk—In looking forward, there is potential interest in seeing the BPP initiative evolve to also address risk as an implicit cost driver. This might allow for innovative ways and means to reduce emerging risks through accelerated or “look ahead” system acquisition and development approaches that emphasize increased reliance on rapid prototyping.

Increase the use of prototyping and experimentation—The rapid development of Web-based or mobile app mashups might be performed by appropriately trained end-users or warfighters (Agre et al., 2014; Endres-Niggemeyer, 2013a). A move towards OA systems for Web-based and mobile devices that rely on apps/widgets retrieved from online marketplaces—apps composed through interpretive software program “scripting” and mashup techniques—is a clear example of this (Endres-Niggemeyer, 2013a; George et al., 2013; Guertin & Womble, 2012; Scacchi & Alspaugh, 2013a). Thus, it is not surprising to find such emerging techniques being investigated and assessed for possible production of new C2 capabilities (George, Bowers, et al., 2014; George et al., 2013; Scacchi & Alspaugh, 2013b).

Achieve dominant capabilities through innovation and technical excellence—An overall summary of the current BBP initiative is focusing attention on how to make acquisition more agile, more innovative, and to develop a new generation acquisition workforce that can enact acquisition processes that are technically excellent—thin and flexible when needed, yet robust and cost-effective, while also being amenable to continuous improvement. This is indeed a real challenge to fulfill, and beyond the scope of what current
acquisition practices are likely to achieve without targeted investment in acquisition improvement research. To be clear, one just needs to consider emerging opportunities (and potential asymmetric cybersecurity threats) that arise through the desire to develop next-generation AC-C2 that are to be composed from apps/widgets that can operate on Web-based/mobile devices. What are the best processes or practices for acquiring, developing, and sustaining deployed systems that are to be built using these new software technologies (e.g., apps/widgets for mobile devices)? How should these processes and practices be adapted to accommodate personal devices (e.g., Apple iPhones/iPads, Android phones/tablets, Blackberry 10 phones) that individual warfighters, joint force troops, or contracted service providers bring with them into the battlespace? How must acquisition processes be best adapted to accommodate and rely on software supply chains that arise around consumer-oriented app marketplaces as possible ways/means for doing more (e.g., rapidly prototyping warfighter composable C2 app/widget mashups [George et al., 2013]) without more (e.g., warfighters who bring their own mobile computing devices for use in C2 contexts; Agre et al., 2014; George, Bowers, et al., 2014)? Once again, these are critical questions to address and resolve through new acquisition research and supporting technology development.

**Emerging Challenges in Achieving BBP Through OA Software Systems for Web-Based and Mobile Devices**

The business models and IP licenses for software components are tightly coupled: Software component licenses codify component producer business models. Said more simply, licenses codify business models. So different software business models imply different software license obligations and rights, and different license types reflect different possible business models. Licenses are generally recognized as contracts regarding IP expressed through terms and conditions that specify obligations and rights stipulated by the component’s producer to enable/constrain what can be done with the component by its integrator or end-users. Understanding and assuring software IP obligations and rights is routinely a task for acquisition offices, and thus a task to be competently performed by the acquisition workforce.

Obligations (like purchase costs/fees paid, or to ensure access to open source software code modifications) denote conditions, events, or actions imposed by a software producer (the licensor) that must be fulfilled by the software integrator/customer enterprise (the licensee) in order to realize the rights identified in the licenses (right to use, right to distribute copies, no right to distribute modified copies, etc.). Note that software system integrators play a role in shaping the obligations and rights imposed on customer enterprises based on choices they make in how software component-based systems are designed, built, and deployed. So where system integration occurs and who does it matters, as does whether customer enterprises that acquire systems have policies that determine which software licenses (or business models) they will accept.

Similarly, we note that “cybersecurity requirements” can also be expressed and analyzed in terms of obligations and rights (Scacchi & Alspaugh, 2011, 2012b). This suggests that the problems and solutions to software component IP license management will be similar in kind or form to those for cybersecurity assurance. Below, we just focus attention on software IP obligations and rights, though the same consequences may apply to the cybersecurity of OA systems and components.
There are many unstated consequences that can arise when software licenses are not well understood. Here are some examples we have seen within the DoD context:

- **Acquisition program managers/staff (including in-house legal counsel) may not understand how software licenses affect OA system design, and vice-versa.** Component-based system design can determine which software licenses will fit, or which can fit if the system design is altered to encapsulate desirable software components with somewhat problematic license obligations or rights (Scacchi & Alspaugh, 2013a).

- **Software license obligations and rights propagate through system development life cycle activities in ways not well understood by system developers, integrators, end-users, or acquisition managers.** We have investigated and described many examples of this in a recent paper that shows how license constraints are mediated by software system design, build-integration, deployment, post-deployment support tools and activities.

- **Different acquisition programs within the DoD and other government agencies may independently reinterpret software component licenses.** This realizes enterprise-wide inefficiencies, as well as increases avoidable costs. It appears to be technically possible to codify software component licenses by type or producer, especially with regards to performative obligations and operational rights that program offices or customer organizations seek. The license modeling techniques we have investigated demonstrate the potential, practicality, and scalability of such possibility (Alspaugh et al., 2012; Scacchi & Alspaugh, 2012a, 2012b, 2013b). However, it may be most efficient and most effective for the DoD to have common legal interpretations for different licenses (or different business models). Such interpretations could be common, if produced by a central legal authority (e.g., Office of General Counsel). Alternatively, it may also be possible for the DoD and other government agencies to provide an open framework or (acquisition) policy guidance whose purpose is to encourage software producers to not only provide software licenses in current narrative forms, but also to provide them in computer processable forms (using domain-specific languages) amenable to automated license analysis. Once again, this is a form of guidance and training we can provide, but it is not one that we can impose on anyone. We believe it is in the best interest of the DoD and other government agencies to employ software licenses that are both human readable and formally processable though automated means, at least in terms of software license obligation and right determinations.

- **Failures to understand software license obligation and rights propagation can reduce DoD buying power, increase software life cycle costs, and reduce competition.** Guidance from the OUSD for Acquisition, Technology, and Logistics recommends programmatic adoption of different BBP 3.0 initiatives grouped into eight focus areas of relevance as methods for innovation, continuous improvements, and doing more without spending more. Acquiring licensed software components is a cost-generating activity, whose costs/fees can be reduced while acquiring ever more agile and adaptive software components and open architecture component-based systems. However, software license non-compliance or worse, infringement, on the part of the DoD will generate costs, cause program delays, as well as reduce agility and adaptation, all of which can be avoided. Such situations can and must be
avoided through acquisition and development practices with little/no additional cost to affect. Such practices can be codified within open source business processes or open source computational business process models that can be shared, customized to specific program needs, redistributed, and archived (Scacchi & Alspaugh, 2013b).

- **Software producers often provide idiosyncratic licenses that generally conform to common business models and common license types.** This seems mainly to arise from efforts by software producers to protect or update their business models in ways that improve their financial yield or protect/lock-in their customer base. This in turn generates demand for time, attention, and effort from legal counsel that support acquisition programs, while also reducing the effectiveness and timeliness of program acquisition efforts. The DoD and other government agencies may be able to explicitly specify in advance what kinds of generic software license obligations they will accept and what kinds of generic software rights they seek, through their own explicit business models. Such specifications can be codified and provided to software producers in open source manner through software license acquisition policies. Software producers might then separate license terms and conditions that do and do not address current license acquisition policies, in order to streamline licensing design and analysis practices for the mutual benefit of software producers, integrators, and customers.

- **Software producers generally provide software licenses that are assumed to legally dominate in systems composed of components from different software producers or integrators.** We refer to software systems (or systems of systems) composed from components (e.g., apps, widgets) subject to different licenses as “heterogeneously-licensed systems” (HLS; Alspaugh et al., 2010; Alspaugh et al., 2012). Popular Web browsers that are compatible with widgets, apps, or plug-in components (e.g., Google Chrome, Mozilla Firefox, Apple Safari) are subject to dozens of component licenses. Popular commercial off-the-shelf (COTS) software components also sometimes encompass components subject to multiple licenses. In both situations, the component producer asserts overall component license obligations and rights in ways that are compatible with the licenses included therein (or so we hope). But when we deploy components that are composed into complex system architectures, or employ components that support on-demand download and implicit integration of smaller components (widgets, plug-ins, scripts, etc.) from online stores, then analysis of license obligation and rights propagation or encapsulation matters. Such technical details can readily overwhelm program acquisition managers and legal staff, thereby reducing the agility and adaptation of component-based system development/deployment. Provision of automated license analysis capabilities within software license management systems should be able to overcome this situation.

- **Given the challenges of HLS, it is unclear what kinds of trade-offs can/should software system integrators or program acquisition staff make in order to maximize overall system development agility and evolutionary adaptation address.** This situation is not unique to the DoD, but is in fact widespread. However, as the DoD and other government agencies move to embrace agile and adaptive component-based software systems to realize new, more timely system capabilities at a lower cost compared to legacy approaches, then
there is a need to provide guidance for how to identify and manage such trade-offs. Failure to recognize the challenges of analyzing and managing HLS systems translates into opportunities lost while avoidable costs increase. We can and should do better than this. But this will require that resources be allocated to identify, articulate, train, and iteratively refine best practices about how, where, when, and why these trade-offs arise. Such knowledge should therefore be captured, codified, shared, accessed, updated, and redistributed in an open source manner.

- **Software IP license and cybersecurity obligations and rights must be tracked, accounted, and managed.** A move to component-based open architecture systems increases organizational overhead for managing software licenses. This overhead can be reduced, or better transformed into productive, value-adding business practices, through the use of automated software obligations and rights management systems (SORMS). While SORMS exist and are routinely used by software component producers (to keep track of who has a licensed copy of their software products), SORMS do not exist at this time for software system integrators or customer enterprises.

- **The DoD and other government agencies would financially and administratively benefit from engaging the development and deployment of an open source automated SORMS.** This may represent the lowest cost means for simplifying license analysis while maximizing the benefits of agile and adaptive component-based software systems acquisition within the DoD and other government agencies. SORMS can help to better DoD software buying power. Similarly, an open source SORMS would also be of value to smaller or startup software producers who may best be able to create innovative and agile software components (widgets) in cost-competitive ways. Last, an open source SORMS intended for software integrator/customer enterprises would be of value to large, established DoD software producers, as a medium through which larger-scale software component acquisitions (e.g., components acquired for standardized deployment throughout an enterprise) can be negotiated and simplified.

- **How best to cultivate and sustain DoD online storefronts and software ecosystem.** The acquisition of development of some DoD Web/mobile widgets may be strongly influenced by commercially available apps that are not secure, nor DoD information assured. Warfighters and others are often drawn to the best available technologies, including apps found in commercial online stores. Who decides whether apps in these conditions should be migrated, secured, and assured to meet DoD requirements? Alternatively, allowing such apps to be used as widgets for rapid prototyping new DoD AC-C2 may represent a promising new direction to stimulate innovation. Subsequently, this entails the needs to better understand possible commercial–DoD online storefront interactions and interdependencies, as well as articulating the needs of DoD agency/program office–specific storefronts. Next, we expect to see redundant app offerings across multiple storefronts, including challenges of identifying common apps of different versions or variants across storefronts and user devices (e.g., is Google Maps the same version across all platforms in use; is Apple Maps equivalent to Google Maps; is Google Earth compatible with NASA World Wind?). How best to determine when redundancy is good/bad for such apps/widgets is unclear and under-explored at this time. Last, as noted, software component
apps/widget licenses and business models across the DoD Software App Ecosystem are very diverse with unclear/unknown interactions and interdependencies. Business models are codified in Web/mobile app IP licenses (e.g., conferring right to use or EULAs) and cybersecurity requirements. Again, much remains here to investigate and resolve to best enable BBP 3.0 initiatives realized with Web-based and mobile software.

Finally, as suggested along the way, all of these consequences can be both anticipated and mitigated through action and careful investment that best enable BBP 3.0 compatible solutions.

New Practices to Realize Cost-Effective Acquisition of OA Software Systems for Web-Based and Mobile Devices

The trends and concerns identified above point to substantial challenges in identifying what can be done to both realize cost-effective BBP for Web-based and mobile device software apps, and to do so in ways that enable and empower the acquisition workforce in the years ahead. Technology, better buying practices, new business models, and new cybersecurity requirements all point to the need for future research and development of new acquisition support technologies, work processes, and guidance practices. The goal is to make sure that acquisition time and effort does not become the main cost and the main risk factor going forward on the path to agile OA Web-based or mobile compatible C2 system development, deployment, and sustaining system evolution.

At this point, we see at least three key areas of opportunity for future acquisition research and development. First, we need to research and develop worked examples of well-formed OA system architectures that are appropriate for C2 system capabilities, and that accommodate Web-based apps, widgets, and mobile devices. Such OA system architectures should specify representative and standardized component interfaces. The examples should also include carefully specified shared agreements that account for different IP licenses and diverse business models of software producers, system integrators, and multiple end-user organizations who must collectively act in ways that enable agile development and adaptive evolution of demonstrable C2 system capabilities.

Second, we need robust open source models of application security processes and reusable cybersecurity requirements that account for exigencies in heterogeneous app/widget software ecosystems, account for software evolution dynamics, formation and continuous improvement of automation-compatible shared agreements, and more. These models should account for description of current process practices, prescription of required verification and validation activities and outcome (deliverable documents or online artifacts), and proscription of what tools/techniques to use, by whom, when, where, and how.

Third, we need precise domain specific languages (DSLs) for specifying, and automated analysis tools for continuously assessing and continuously improving, cybersecurity and IP license requirements for dynamically evolving Web/mobile C2 system-based capabilities. The DSLs needed must be able to specify and operationalize the shared agreements between different DoD organizations, government agencies, and commercial enterprises involved in producing, integrating, or evolving component-based OA C2 system capabilities.

Overall, what we call for is similar in kind to what we have already produced and applied in other software development domains, using then current technologies (Jensen & Scacchi, 2005; Scacchi & Alspaugh, 2008). What we now call for is a reinvention and repurposing of these concepts, but in contemporary forms scaled and secured in ways that
best meet the needs of the DoD program offices, acquisition program managers, and others in the acquisition workforce to best support BBP 3.0 initiatives for Web-based and mobile device software components (widgets, apps, plug-ins).

**Conclusions**

The DoD, other government agencies, and most large-scale business enterprises continually seek new ways to improve the functional capabilities of their software-intensive systems. The acquisition of OA systems that can adapt and evolve through replacement of functionally similar software component applications (apps) and widgets is an innovation that can lead to lower cost systems through more agile system development and adaptive system evolution. Our research identifies and analyzes how new software component apps and widgets, their IP license and cybersecurity requirements, and new software business models can interact to drive down (or drive up) total system costs across the system acquisition life cycle. The availability of such new scientific knowledge and technological practices can give rise to more effective expenditures of public funds and improve the effectiveness of future software-intensive systems used in government and industry.

Our study reported in this paper also identifies a new set of technical risks that can dilute the cost-effectiveness of Better Buying Power efforts. It similarly suggests that current acquisition practices aligned with BBP can also give rise to acquisition management activities that can dominate and overwhelm the costs of OA system development. This adverse condition can arise through app/widget vetting, new software business models, opaque and/or underspecified acquisition management processes, and the evolving interactions of new software development and deployment techniques. Unless proactive investment in acquisition research and development can give rise to worked examples, open source models, and new acquisition management system technologies, the likelihood of acquisition management dominating agile development and adaptive deployment of component-based OA C2 system capabilities.

Overall, this paper serves to help describe and detail how Web-based and mobile device software component technologies, IP licenses, security requirements, business models, and adaptive system evolution interact. It also highlights what policies, practices, or technologies within the DoD and other government agencies can simplify or exacerbate OA system cost arising at different points in the acquisition life cycle. Our common goal is to increase the ways, means, and beneficial consequences of the transition to the cost-effective acquisition of Web-based and mobile device OA software systems whose acquisition, development, deployment, and ongoing evolution are agile and adaptive.

**References**


Acknowledgements
The research described in this report was supported by grants N00244-14-1-0030 and N00244-15-1-0010 from the Acquisition Research Program at the Naval Postgraduate School, Monterey, CA. No endorsement, review, or approval implied. This paper reflects the views and opinions of the authors, and not necessarily the views or positions of any other persons, group, enterprise, or government agency.
**Introduction**

From remotely piloted aircraft and smart bombs to autonomous vehicles and advanced fighter jets, software is crucial to the success of today’s weapon systems. Focusing solely on developing and maintaining military hardware is no longer an option. With shrinking defense budgets and increasingly complex systems, the defense industry and services must fight to deliver on this ambitious objective, the military must drastically transform its approach to software. New organizational structures, operating models, and tools will be essential to modernizing and sustaining the U.S. weapon systems. (Hagen, Hurt, & Sorenson, 2013, p. 31)

Although the Department of Defense (DoD) has developed some very successful software-intensive systems, such as the Aegis, Tomahawk Missile, and F/A-18 Hornet, we continue to struggle with successfully developing like systems. The software development in the F-35 Joint Strike Fighter (JSF) continues to be problematic. The GAO (2012) stated that

JSF software development is one of the largest and most complex projects in DoD history, providing essential capability, but software has grown in size and complexity, and is taking longer to complete than expected. Developing, testing, and integrating software, mission systems, and logistics systems are critical for demonstrating the operational effectiveness and suitability of a fully integrated, capable aircraft and pose significant technical risks moving forward. (p. 7)

The report went on to state, “This program [JSF] has modified the software development and integration schedule several times, in each instance lengthening the time needed to complete work” (GAO, 2012, p. 11). The results of the software development problems have contributed to a two-year delay and increased costs of about one billion dollars.

When software-intensive systems encounter developmental problems, it is easy to see the symptoms: schedule overruns, acquisition cost overruns, systems delivered with less capability than desired, and unaffordable software sustainment costs. The actual causes of the visible symptoms are often much more difficult to determine.

Cost and schedule overruns in software development are often the result of poor initial software size estimates and unforeseen software redesign. In the case of the JSF,
The lines of code necessary for the JSF’s capabilities have now grown to over 24 million—9.5 million on board the aircraft. By comparison, JSF has about 3 times more on-board software lines of code than the F-22A Raptor and 6 times more than the F/A-18 E/F Super Hornet. This has added work and increased the overall complexity of the effort. The software on-board the aircraft and needed for operations has grown 37 percent since the critical design review in 2005 … almost half of the on-board software has yet to complete integration and test—typically the most challenging phase of software development. (GAO, 2012, p. 11)

The report goes on to state that typical software size growth in DoD systems development ranges from 30% to 100%.

JSF design changes were originally supposed to taper off and be completed by January 2014. Actual design changes through September 2011 failed to taper off and continue at a significantly high rate. The projections in the GAO (2012) report indicated that the revised design change projections will continue, and actually grow in number, until January 2019 (p. 16). Given this level of redesign, the software and system complexity growth are likely to continue.

The DoD Software-Intensive System Development Problem and Research Technique

Problem

From a systems management perspective, the overarching problem is that the DoD Acquisition Management System produces both successful and unsuccessful software-intensive systems. The management oversight, structure, and discipline offered do not produce repeatable success in complex, software-intensive systems development.

Primary Research Question

The problem previously identified drives this primary research question: Why does the DoD Acquisition Management System produce both successful and unsuccessful software-intensive systems?

Secondary Research Questions

I analyze the DoD software-intensive system development challenge by addressing these secondary research questions:

- Does the DoD acquisition environment provide opportunity for variable results in software-intensive system development?
- How does the software engineering environment impact DoD software intensive system development?
- Is the DoD requirements development and communication process sufficient for potential software developers?
- How is the software-intensive system architecture developed to ensure warfighter capabilities are designed and prioritized?

DoD Acquisition Environment

At the top level, there are the three primary decision support systems used within the DoD, and the interaction within these systems significantly decides the acquisition of products or services (DoD, 2013b). The three systems are the Joint Capabilities Integration and Development System (JCIDS), which provides the acquisition requirements documents;
the Defense Acquisition System (DAS), which provides the processes to develop and acquire the needed products to fulfill the requirement; and the Planning, Programming, Budgeting, and Execution (PPBE) process, which is the funding resource management.

Software-intensive systems are most impacted by the JCIDS and the DAS Decision Support Systems, and the PPBE process has no particularly unique impact on software intensive systems development. This research, therefore, focuses on elements of the JCIDS and DAS systems.

Requirements Generation

The Joint Capabilities Integration and Development System (JCIDS) was designed to assess capability requirements and associated capability gaps and risks (CJCS, 2012, p. A-1). Capability gaps may be identified in one or more of the following areas: Doctrine, Organization, Training, Materiel, Leadership Policy and Education, Personnel, Facilities, and Policy (DOTMLPF-P). Materiel-related capability gaps become the basis for the requirements process that drives the acquisition community to develop and acquire platforms designed to bridge all or part of the identified gap. JCIDS is designed to be an iterative process, beginning with a validated Initial capabilities Document (ICD), triggering the acquisition community to begin an Analysis of Alternatives (AoA) on candidate systems that potentially address the capability need. The Capabilities Design Document (CDD) refines and adds necessary detail to support the technical design of the system sought. The final document in the series is the Capabilities Production Document (CPD), which further refines the user requirements and adds detail supporting the production planning for the system. Although JCIDS is designed to refine well-defined requirements, there is clearly an opportunity for requirements creep with this iterative user requirements process.

After the user community completes each JCIDS iteration, the program/project/product manager (PM) or materiel developer is prompted into action. As stated, the ICD prompts an AoA identifying the possible systems that could be procured or developed to meet the capability need. The CDD is a key document in the requirements generation cycle and is the user community’s primary input for the PM’s development of the performance specification for the Request for Proposal (RFP). The CPD is the user’s key document for driving production decisions, and the PM’s production strategy is significantly influenced by the CPD.

One of the PM’s most critical functions is developing the performance specification for inclusion in the RFP. This requires the PM team to translate the user-stated needs from capabilities-based language to performance-based language that is used to drive the design efforts of potential system developers, usually contractors. This is critical because the RFP is the basis for the potential contractors’ proposals containing the estimated cost, schedule, and technical performance they plan to achieve. The submitted proposals are evaluated and compared during the labor-intensive source selection process, resulting in a contract award based on proposal merit. If the performance specification is incomplete, vaguely stated, or misunderstood, then the source selection process and contract award is based on incorrect proposals and the effort is significantly wasted.

The selected contractor accepts the terms of the contract based on the assumptions and estimates contained in the proposal. To develop the proposal, the contractor translates the PM’s performance specification into a basic detailed specification so that the scope of work can be estimated for the proposed cost and schedule. Correcting these performance specification deficiencies later puts the government at a significant disadvantage, as the contract has been awarded and necessary changes to the contract are negotiated without competition. Changes, additions, or even clarifications to the performance specification after
contract award are likely to impact the terms of the contract, resulting in a negative impact to the cost, schedule, or performance of the desired system.

**The Defense Acquisition System**

The DoD Acquisition, Technology, and Logistics Life Cycle Management System is the framework for control and management of DoD systems development, based on the SEP. The model features development phases that define activities, and milestones that serve as control and decision points. These phases and milestones are established very early in the development cycle using the information available during early Materiel Solution Analysis (MSA), which is obviously very limited. Overwhelmingly, the PM responsible for establishing this strategy is not the individual responsible for executing it. Funding requirements, including amount, type, and *period of execution*, are established in the Program Objective Memorandum (POM), submission and a congressionally-approved funding profile is established for the entire acquisition strategy within the PPBE process. At this point, the schedule becomes very rigid as Congress must approve significant changes to the funding profile, including when the funding is to be executed. Although there are obviously known and unknown risks associated with an acquisition strategy formulated this early, there is no provision for a management reserve of funding to address these risks.

The Interim DoD Instruction 5000.02, dated November 26, 2013, shows alternate versions of the DAS phases and milestones (see Figure 3) that attempt to address the impact that software imparts on the development process. The interim instruction depicts the following variants of the model: Defense Unique Software Intensive Program; Incrementally Fielded Software Intensive Program; Hybrid Program A (Hardware dominant); and Hybrid Program B (Software Dominant) models (DoD, 2013b, pp. 10–14). The new models indicate an understanding that software impacts the system development process differently than typical hardware systems do. As these are all newly developed, their impact on future development is unknown.

**Performance Specifications and the Work Breakdown Structure**

Since the implementation of acquisition reform in the nineties, detailed specifications have been replaced with performance specifications in order to leverage the considerable experience and expertise available in the defense contractor base. In most hardware-centric engineering disciplines, the expertise that the DoD seeks to leverage includes a mature engineering environment in which materials, standards, tools, techniques, and processes are widely accepted and implemented by industry leaders. This engineering maturity helps to account for derived and implied requirements not explicitly stated in the performance specification. Three levels of the work breakdown structure (WBS) may provide sufficient detail for vendors to develop a desired system in a mature engineering environment, such as the automotive field. For example, an automotive design that provides for easy replacement of wear-out items such as tires, filters, belts, and batteries obviously provides sustainability performance that is absolutely required. Most performance specifications do not explicitly address this capability as they would be automatically considered by any competent provider within the mature automotive engineering environment.

The *Department of Defense Handbook: Work Breakdown Structures for Defense Materiel Items* (MIL-HDBK-881A), recommends a minimum of three levels be developed before handoff to a contractor (DoD, 2005). If a program is expected to be high cost or high risk, it is critical to define the system at a lower level of the WBS (DoD, 2005, p. 3). Complex weapon systems are nearly always high cost, and the complex software development that these systems require almost always means that the development effort is high risk as well. The WBS and performance specification must, consequently, be significantly more
developed to provide the software engineer enough information and insight to accurately estimate the level of effort needed—cost and schedule—and to actually produce the capabilities needed by the warfighter. Contracts resulting from proposals that are based on underdeveloped, vague, or missing requirements typically result in catastrophic cost and schedule growth as the true demands of the software development effort are discovered only after contract award.

**Technology Readiness Assessment and Risk Management**

Another important management aspect is addressing the readiness of the key technologies for successful development and deployment. A Technology Readiness Assessment (TRA) is required for most Major DoD Acquisition Programs (MDAPs; DoD, 2011, p. 1-1) The purpose for conducting a TRA is to address the risk of attempting to develop a system with a key technology that is too immature to successfully deploy the system when needed by the warfighter. To benchmark the assessment, Technology Readiness Levels (TRLs) have been developed in a nine-level model, with a goal of ensuring that a system’s key technologies achieve at least a TRL level 6 to reduce risk to an acceptable level.

There are software TRLs established, and level 6 is defined as “Module and/or subsystem validation in a relevant end-to-end environment.” The level 6 description specifies the “level at which the engineering feasibility of a software technology is demonstrated. This level extends the laboratory prototype implementations on full-scale realistic problems in which the software technology is partially integrated with existing hardware/software systems” (Blanchette, Albert, & Garcia-Miller, 2010, p. 35).

The software TRL level 6 description presents several problems in performing the TRA on a software-intensive system. Weapon system software is typically engineered from scratch with few reused elements, which means that there is very little-to-nothing on which to perform the assessment. There will likely be software developed for similar systems that would meet the level 6 description, but assessing like-software built for another system will not significantly reduce the software technology risk of the proposed system. For example, the F-35 is built by the same manufacturer as the F-22, and they are both high-performance military aircraft with different but overlapping missions. Yet the F-35 is experiencing more software development problems than its predecessor and already has three times more software than the F-22 (Hagen et al., 2013, p. 26).

Software TRLs do not appear to be providing the same type readiness indicator as hardware-related TRLs, leaving software technology risks substantially unknown. In a 2010 U.S. Army workshop report from the Software Engineering Institute (SEI), the participants noted that “though marginally useful, these efforts have only confirmed for the participants the futility of continuing to base [technology] readiness decisions for software aspects of systems on the DoD software TRLs” (Blanchette et al., 2010, p. 2). The software TRLs clearly do not seem to be effective at reducing risk for the TRA.

To help with early risk management in lieu of effective software TRLs, a software developer maturity assessment is mandated for most software-intensive systems, through attaining level 3 in the SEI’s Capability Maturity Model Integrated (CMMI), or equivalent, assessment methodology (DoD, 2013a, p. 92) The concept recognizes that the software build is a product of the process, and more mature organizations—those with successful past performance, demonstrated engineering discipline, stable development staffs, and effective management structures—reduce system development risk.

SEI also has the Software Acquisition Capability Maturity Model® (SA-CMM), which is designed to evaluate the maturity of software acquiring organizations such as the DoD’s
software-intensive system PM offices (Cooper & Fisher, 2002). The SA-CMM is also a five-level model, similar to the CMMI. The DoD currently has no requirement for PM offices to undergo an evaluation or achieve any SA-CMM level, but the maturity of the team responsible for communicating the system requirements and managing the development has an impact on risk.

**Findings Summary**

In summary, the DoD acquisition environment features a requirements flow-down process that involves user-stated capabilities-based requirements translated to performance-based requirements, then translated to the detailed design specifications. This requirements translation process is the basis for the resource-intensive source selection and binding contracting processes, which are critical for accurate cost and schedule estimates. Although DoD acquisition is based on the event-driven SEP, the schedule becomes rigid very early in the process when time-specific funding is attached. The subsequent system PMs are charged with managing the cost, schedule, and performance set by the initial PM with no funding provided for managing the associated risk. To reduce risk, PMs are directed to perform TRAs early in the process, with a goal of achieving at least TRL 6 on key technologies. Software TRLs do not appear to be effective, and software developer maturity assessments are conducted to help reduce system development risk. The latest Interim DoD Instruction 5000.02 (DoD, 2013a) depicts newer phases and milestone models that attempt to address the differences that software development causes in the management of the DAS.

**DoD Acquisition Environment Analysis**

Does the DoD acquisition environment provide opportunity for variable results in software-intensive system development?

The DoD acquisition environment appears to remain vulnerable to significant variability when developing software-intensive systems, similar to the problems currently plaguing the F-35 JSF program. Although the new phases and milestones models address the software component development, other critical management functions remain unchanged. Requirements generation, performance specification development, RFP, source selection, and contracting processes have yet to adapt to the unique challenges presented when managing software-intensive system development. Early program risk management assesses key technology readiness, but the software TRLs are ineffective for predicting software development risk. Evaluating the software developer’s maturity helps reduce some risk but fails to include the critical DoD entities in any maturity assessment.

Early risk management through the TRA and achieving a desired TRL is ineffective for the software component. Assessing the contractor (software developer) maturity through CMMI or equivalent evaluation appears to be effective in reducing the developer risk but does not address the DoD acquisition community maturity. As the software developer is significantly dependent on the government’s ability to effectively generate and clearly communicate a comprehensive set of requirements, quality attributes, and critical design elements, assessing just the developer’s maturity addresses only part of the risk.

**Software Engineering Environment**

**Software Engineering**

The software engineering environment is not mature, especially when compared to hardware-centric engineering environments. Dr. Philippe Kruchten (2005) of the University of British Columbia remarked, “We haven’t found the fundamental laws of software that would play the role that the fundamental laws of physics play for other engineering
disciplines” (p. 17). Software engineering is significantly unbounded because there are no physical laws that help define environments. There is significant evidence for software engineering immaturity, and it is nearly impossible to find widely accepted, industry-wide development standards, protocols, architectures, or formats. There is no dominant programming language, design and development process, standard architectures, or software engineering tools, which means that reusable modules and components rapidly become obsolete. All of these combine to make it nearly impossible to institute a widely accepted software reuse repository. Without significant software architecture and code reuse in developing software-intensive weapon systems, each development process essentially starts from scratch. This fact is one of the main reasons that the TRA and the software TRLs are ineffective in predicting software development risk (Naegle & Petross, 2007).

The software engineering state-of-the-practice currently is wholly dependent on the requirements that are passed to the software development team. From the requirements, a software architecture is designed, and the requirements “flow down” through that architecture to the individual modules and computer software units that are to be constructed. The software build focuses on the requirements that flowed down to that level and the integration required for functionality. The standards, protocols, formats, languages, and tools used for the build will likely be unique to the contractor developing the software, and will most certainly not be universally accepted or recognized across the software industry.

The software architectural design is the basis for all of the current and future system performance that the system will achieve, and the current state-of-the-practice in software engineering has each project design a unique architecture. Like hardware, the software design will significantly impact system attributes that are important to the warfighter, including maintainability, upgradability, interoperability, reliability, safety, and security. Most hardware-oriented engineering environments address these critical areas through widely accepted industry standards.

**Findings Summary**

With software, virtually all of the performance and quality attributes developed come directly from the requirements received, and the immature software engineering environment will likely not compensate for any desired performance, such as system sustainability, that is not clearly specified in a requirement. Unlike hardware-oriented engineering environments, where the widely accepted industry standards will be employed whether or not they are specified, with software, you get what you specify and very little else.

The software architectural designs suffer from the immature engineering environment as well. Each software design is unique and driven by the requirements received with no industry-standard architectures available. All current and future system attributes impacted by the architecture must be communicated to the software design staff to ensure they are considered in the design process.

**Software Engineering Environment Analysis**

How does the software engineering environment impact DoD software-intensive system development?

As illustrated in the previous section, the lack of software engineering maturity impacts both requirements development and design of the architecture. To compensate for the relative immaturity of the software engineering environment, the DoD must conduct
significantly more in-depth requirements analysis and provide potential software developers
detailed performance specifications in all areas of software performance and sustainability.
This is a significantly different mind-set than the hardware-dominated systems acquisition of
the past.

In addition to the performance requirements, software architectures must be similarly
shaped to include system attributes expected by the warfighter. Many DoD user
representatives and acquisition professionals have grown accustomed to the engineering
maturity levels offered by the hardware-oriented systems that dominated past acquisitions.
Providing the system requirements in the same fashion may not drive the architecture for
needed attributes. As demonstrated by the F-35 JSF redesign problems, changing software
architectures during the development cycle will likely be costly in terms of schedule and
funding.

**DoD Acquisition Environment: Impact on Software Development and Quality Attributes**

*DoD Requirements Generation Process*

The DoD requirements generation process was described earlier as part of the DoD
acquisition environment and consists of three major processes: user-generated
requirements in the form of *capability needs* using the JCIDS; PM-generated requirements
in the form of *performance specifications*; and finally, contractor-generated *detailed
specifications*, developed generally in that order. Two major requirements language
interpretations are required to get from the warfighters’ needs to the system built to meet
those needs, leaving significant opportunity for misinterpretation, omission, and
misunderstanding of weakly articulated and vaguely stated language. To do this effectively,
the PM must accurately interpret user capability language (for example, warfighter requires
the capability to … in all mission environments) and translate that into performance
language (for example, system shall achieve xxx performance … in these specific
conditions). The contractor then translates the performance language into the system build-
details that meet or exceed the performance specified.

The importance of system software requirements development to the potential
success of software-intensive systems development cannot be overstated. Underdeveloped,
vaguely articulated, ill-defined software requirements elicitation has been linked to poor cost
and schedule estimations, resulting in disastrous cost and schedule overruns such as what
the F-35 JSF is currently experiencing. In addition, the resulting products have been lacking
important functionality, are unreliable, and have been costly and difficult to effectively
sustain (Naegle, 2006).

*Systems Engineering Process*

Using the SEP approach, the explicit user capabilities requirements specified in the
Joint Capabilities Integration and Development System (JCIDS) provides the input for
system requirements analyses. These analyses are intended to illuminate all system-stated,
-derived, and -implied requirements and quality attributes necessary to achieve the
capabilities needed by the warfighter. The WBS is a methodology for defining ever-
increasing levels of performance specificity using the SEP to guide the development of each
successive layer (DoD, 2005, pp. 1–5).

Just as it supports hardware development, the Systems Engineering Process
(SEP) is essential in the development of software design. In software
development, good quality and predictable results are paramount goals in
creating the specified warfighter capabilities within cost and schedule
constraints. To accomplish those goals, we examine the methods, tools and processes the software developer uses in building the software with the intent of attaining a product that provides all of the necessary functionality and is supportable, efficient, reliable and easy to upgrade. (Naegle & Petross, 2007, pp. 14, 15)

**Software Engineering Maturity Impact on Requirements Generation**

The immature software engineering environment, discussed earlier, can be compensated for only by a requirements generation system that does not leave any gaps in performance or quality attributes needed. Having all of the requirements clearly communicated is critical, but the software engineer must also understand the requirements in context. Both essential and enhancing features are communicated to the system and software developers as requirements and, as such, appear to have equal weight. The critical difference between “essential” and “enhancing” may not be clear to the software development team, which may result in a poorly performing and possibly dangerous design. The distinction needs to be made clear, but there is no definitive method for identifying requirements as system “essential” or “enhancing.”

**System Operational Context**

To gain some insight into the operational environments that the system is expected to operate within, the DoD provides an Operational Mode Summary/Mission Profile (OMS/MP). The OMS/MP provides some basic insight into the operational profile, threat profile, environmental profile, and the terrain/sea state/undersea/air environment profile, which adds some context to the requirements, but is not usually scenario based. It typically lacks sustainability activities, interoperability profiles, system life-cycle profiles, planned or anticipated upgrades, or operation in stressful, degraded, or emergency situations. There is no prioritization of the operational modes or configurations, nor identification of critical and non-critical systems.

The software development team would likely continue to be missing important information that it needs to adequately design the software and to predict the funding and schedule resources necessary to build the software the warfighter expects.

The OMS/MP documents do not typically provide any information regarding system life-cycle changes such as pre-planned product improvement (P3I) programs, planned upgrades and technology refreshments, future interoperability requirements, or plans for future integration into tactical and logistical networks. These life-cycle events, while known or anticipated, are not effectively communicated to potential developers for inclusion in the proposal process and are often omitted from the software system design.

**Impact on Software and Quality Attributes Analysis**

Is the DoD requirements development and communication process sufficient for potential software developers?

The DoD requirements generation process that was purposefully designed to garner the maximum contractor innovation and flexibility appears to provide too little information for the software developer to adequately predict the resources necessary to develop the system software. It is clear that the current state of the software engineering environment is mostly incapable of compensating for missing, vaguely stated, or weakly articulated requirements. At the same time, the current DoD requirements generation system provides ample opportunity to inadvertently omit requirements and to provide vaguely stated or weakly articulated requirements through the capabilities-oriented JCIDS documents and the performance-based specifications derived from them.
Without fully understanding the requirements in a detailed operational context, the software design and development effort and resources remain significantly unknown. The typical OMS/MP provides some operational context to the requirements, but is not sufficiently detailed to provide the design drivers needed by the software engineers. Developing a proposal with this limited information will likely result in a significantly underestimated software development effort. After contract award, more operational details are typically provided through program and design reviews, and the cost and schedule for the software effort are likely to inflate significantly to accommodate the new understanding of the requirement in a non-competitive environment.

The lack of operational context typically provided by the government during the RFP process appears to have significant negative impacts on the software design for reliability and maintainability. The OMS/MP documents’ lack of information regarding significant planned and anticipated life-cycle changes, system sustainment activities and burden, and operations under unusual conditions will likely mean that the system software design will not easily accommodate known changes. There is no prioritization of the operational modes or configurations that would impact system design considerations. This information would also help differentiate critical systems from enhancing (non-critical) systems, providing a priority in the software design effort.

**Software-Intensive System Architecture Development Analysis**

How is the software-intensive system architecture developed to ensure warfighter capabilities are designed and prioritized?

The DoD system architectural process, with all of its tools, techniques, and discipline, appears to be ineffective in driving repeatable, successful software designs. Within the SEP, there are three DoD processes that drive the system architecture: the requirements generations system, the WBS, and the OMS/MP.

There appears to be significant opportunity to omit requirements, or to provide vague or weakly articulated requirements through the translation process from the user capability-based requirements, to the PM’s performance specification, and finally to the contractor’s detailed specification. This problem is exacerbated by the immature software engineering environment described earlier, which is solely focused on requirements as provided.

The process of developing the WBS appears to be similarly flawed in effectively communicating the functional architecture to a sufficient level for the software developers. The overarching philosophy for both requirements generation and the WBS, in order to garner the maximum flexibility and innovation, is purposely not to be specific. Due to the immature engineering environment, the software components need significantly more specificity than the hardware counterparts to produce realism in the cost and schedule provided in the contractor’s proposal.

The operational context information that the government provides appears to be insufficient for the potential software developers to have an understanding of the requirements within the context of the operational environment, constraints, and life-cycle events of the proposed system. The OMS/MP typically provides only a vague understanding of the operational environment and significantly more information is required to design and build the system actually needed by the warfighter. This additional information is likely to be added in program and design reviews conducted after the contract is awarded, so resulting changes impacting the software development can cause significant increases in the cost and schedule, all negotiated without the advantages of a competitive environment.
Conclusions

The DoD acquisition process provides the environment for both successful and unsuccessful software-intensive systems development. Specific elements of the DoD acquisition process that contribute to the variable environment include the following:

- **The DoD Requirements Generation Process.** The translation process from JCIDS capabilities-based language to the RFP/contract performance-based language, and finally to the specification-based detailed language creates ample opportunity for misinterpreted requirements to be communicated. This process was designed to garner innovation from mature engineering fields that leverage widely accepted materials, processes, and standards—attributes that the software engineering field does not yet have.

- **Communicating Operational Context.** The Operational Mode Summary/Mission Profile (OMS/MP) provides some insight into a system’s intended operational context but provides far too little information for the complex software design process. This lack of detail, again, cannot be compensated by the immature software engineering environment and so impacts software-intensive systems more than hardware-centric ones.

- **Failure to Compensate for the Immature Software Engineering Environment.** As demonstrated by the first two bullets, one of the major differences between successful and unsuccessful software-intensive systems development is recognizing and compensating for the immature software environment. The DoD Acquisition System policies, guidelines, and controls do not provide a framework to ensure that essential software attributes are sufficiently revealed and effectively communicated to the contractors that will design and build the software systems.

- **The DoD Acquisition System**
  - The DAS is designed to leverage industry innovation by providing performance specifications that are designed to allow mature industrial engineering environments to develop the best-value technologies that meet the performance specifications. This is effective when the engineering environments are mature and can offer viable, mature technology alternatives that are considered industry standard. There are insufficient DAS processes for recognizing and compensating for immature engineering environments, such as exists in the software field.
  
  - The schedule and funding profile are initially set by the first system PM, and the program depends significantly on how well the requirements generation process accurately identified the bulk of the requirements. Once funding is linked to milestones, the program cost and schedule become very rigid, which exacerbates problems with software-intensive system developments that have late requirements creep due to insufficient understanding of the effort in the proposal preparation.
  
  - Software Technology Readiness Levels (TRLs) are ineffective in reducing risks associated with the system software development. Because there are few reusable software components, limited industry-wide standards for architecture and supportability, and rapidly
emerging languages, protocols and tools, the software TRLs, based on past efforts, are not reliable predictors of software readiness.

- Software development significantly adds to the system development risk. The DAS is designed to reduce development risk, but cannot eliminate all associated risks. Some risk is accepted with the expectation that the PM team will effectively manage those risks, yet there is no funding management reserve provided to do so. Any risk management mitigation effort that involves funding has the opportunity to create a cascade of management actions resulting from funding reductions in other planned and necessary activities.

Recommendations

General

As part of this research, I searched for tools, techniques, and procedures that would address the software-intensive system development problems and integrate well with the Defense Acquisition System (DAS) while supporting the Systems Engineering Process (SEP). The tools, techniques, and procedures recommended in this section are not particularly new and many programs may have used some, most, or all of these in the development of their systems. The major recommendation is that DoD formalize and institute the use of these tools, techniques, and procedures (or similar ones) for the development of software-intensive systems. There would almost certainly be a benefit when applied to hardware-centric system development, too, and certainly there would be no detriment in using them for all complex system development.

One of the findings of this research was the lack of a PM management reserve fund to address accepted development risks, but a significant policy and political change would be required to provide a management reserve in program funding. I believe this course of action to be unlikely, but the implementation of the recommendations would significantly reduce software-intensive system developmental volatility and risk, and reduce the need for the management reserve.

Each of the tools, techniques, and procedures are valuable in assisting the systems development process, but when used together, provide a synergistic effect to the vital front-end analyses that directly impact the shortcomings revealed in this research. Implementing these tools does not require any major adjustments to the DAS or the SEP, and in fact become major enablers for both.

Tools, Techniques, and Processes

The following tools, techniques, and processes are briefly described in this section:

- The Software Engineering Institute’s (SEI’s) Quality Attribute Workshop (QAW)
- The Maintainability, Upgradability, Interoperability, Reliability, & Safety and Security (MUIRS) analytic technique
- The Software Engineering Institute’s Architectural Tradeoff Analysis Methodology (ATAMsm)
- The Failure Modes and Effects Criticality Analysis (FMECA)
- Software Management Readiness Levels (MgtRL)
**Quality Attribute Workshop**

The QAW is primarily a method for more fully developing system software requirements and is intended to provide stakeholders’ input about their needs and expectations from the software (Barbacci et al., 2003, p. 1). As the system requirements are developed, software quality attributes are identified and become the basis for designing the software architecture.

The SEI’s QAW is implemented before the software architecture has been created and is intended to provide stakeholder input about the needs and expectations from the software (Naegle, 2007). The QAW process provides a vehicle for keeping the combat developer and user community involved in the DoD acquisition process, which is a key goal of that process. In addition, the QAW includes scenario-building processes that are essential for the software developer to design the software system architecture (Barbacci et al., 2003, pp. 9–11). These scenarios will continue to be developed and prioritized after contract award to provide context to the quality attribute identified for the system.

**Primary Software Acquisition Problem Area Addressed**

The QAW process is primarily designed to more fully develop system software requirements so that the government RFP is clearer to potential contractors. In turn, the resulting proposals should be more accurate and realistic, reducing requirements and project scope creep.

**Maintainability, Upgradability, Interoperability/Interfaces, Reliability, and Safety/Security Analytic Technique**

The MUIRS analytic technique is designed to provide a framework for better understanding of essential supportability and safety/security aspects that the warfighter needs and expects but often doesn’t communicate clearly with the capabilities-based JCIDS documents. This analytic technique helps compensate for the immature software engineering environment as the MUIRS analysis illuminates the derived and implied requirements that the immature environment cannot.

Much of the software supportability and safety/security performance that typically lacks consideration and is not routinely addressed in the software engineering environment can be captured through development and analysis of the MUIRS elements. Analyzing the warfighter requirements in a QAW framework for performance in each MUIRS area will help stakeholders identify software quality attributes that need to be communicated to potential software contractors (Naegle, 2006, pp. 17–24).

The MUIRS analysis assists the QAW process by focusing on those elements that are too often typically overlooked during the requirements generation process. The QAW and MUIRS analysis are critical to the software design process, discussed in the next section.

**Primary Software Acquisition Problem Area Addressed**

MUIRS primarily addresses the immature software engineering environment as it provides an analytic approach for critical sustainment and safety/security attributes often missing, weakly articulated, or vaguely stated in the requirements produced. With its capabilities and performance-based requirements processes, the DoD significantly depends on mature engineering environments to fill the gaps left from the requirements generation and communication processes, but the software engineering environment is unable to do so. The MUIRS analysis is also an enabler for the QAW and ATAMsm architectural processes discussed next.
Architectural Tradeoff Analysis Methodology

The SEI’s ATAM is an architectural analysis tool designed to evaluate design decisions based on the quality attribute requirements of the system being developed. The methodology is a process for determining whether the quality attributes are achievable by the architecture as it has been conceived before enormous resources have been committed to that design. One of the main goals is to gain insight into how the quality attributes trade off against each other (Kazman, Kleim, & Clements, 2000, p. 1).

Within the SEP, the ATAM provides the critical Requirements Loop process, tracing each requirement or quality attribute to corresponding functions reflected in the software architectural design. Whether ATAM or another analysis technique is used, this critical SEP process must be performed to ensure that functional- or object-oriented designs meet all stated, derived, and implied warfighter requirements. In complex systems development such as weapon systems, half or more than half of the total software development effort is expended in the architectural design process. Therefore, the DoD PMs must ensure that the design is addressing requirements in context and that the resulting architecture has a high probability of producing the specified warfighters’ capabilities described in the JCIDS documents.

The ATAM focuses on quality attribute requirements, so it is critical to have precise characterizations for each. To characterize a quality attribute, the following questions must be answered:

- What are the stimuli to which the architecture must respond?
- What is the measurable or observable manifestation of the quality attribute by which its achievement is judged?
- What are the key architectural decisions that impact achieving the attribute requirement? (Kazman et al., 2000, p. 5)

The ATAM is designed to elicit the data and information needed to adequately address the three previous questions. These questions, focused on requirements and quality attributes, are user-centric, and so the ATAM scenarios must be constructed by the user community (Naegle & Petross, 2007, p. 25). The methodology keys on scenario development in three main areas:

- **Use Case Scenarios.** As the name suggests, these scenarios describe how the system will be used and sustained in the harshest environments envisioned. It includes all interoperability requirements and duty cycles as well.

- **Growth Scenarios.** Growth scenarios focus on known and anticipated system change requirements over the intended life cycle. These scenarios include upgrades and technology refreshments planned; interoperability requirements, such as inclusion in future warfighting networks; changes in sustainment concepts, and other system changes expected to occur over time.

- **Exploratory Scenarios.** Exploratory scenarios focus on operations in unusual or stressful situations. These address user expectations when the system is degraded or operated beyond normal limitations due to emergency created by combat environments. These scenarios include Failure Modes and Effects Criticality Analyses (FMECA) to identify the essential functions that must not fail. As important to the software engineers, FMECA also identifies those enhancing functions that should not preclude the system from
functioning when that enhancing function is degraded or non-operational. The software engineers need that information to properly design the software.

Test cases are developed out of the scenarios, which firmly link the test program with the user requirements in the context of the scenarios. This methodology also helps to ensure that there are verification events for software and sustainment requirements, which are too often missing from the testing program.

As shown in Figure 1, the ATAM is an integrating function for many of the tools and techniques discussed here. It is designed to be an iterative process and would be most effective when started in early concept development, then continued through contract award, prototyping, and into the design review process.

![Figure 1. Quality Attribution Workshop and Architectural Tradeoff Analysis Methodology Integration Into Software Life-Cycle Management (Naegle & Petross, 2007, p. 25)](image)

**Primary Software Acquisition Problem Areas Addressed**

The ATAM process addresses four primary problem areas:

- The scenario development provides much more operational context than the typical OMS/MP provides. This level of detail helps to compensate for the immature software engineering environment and is critical for the proper design of the software architecture.
- The ATAM serves as a very effective software design metric function. With the software development team using 50% or more of the available resources for requirements analysis and software design before the Preliminary Design Review (PDR), it is critical to have an effective software design metrics function. Traditional software design metrics focus on the design complexity...
and do not address whether the design is adequate or not. ATAM directly links the user requirements to the system architectural design.

- As the testing program is developed from the scenarios, it becomes difficult to omit any critical testing event. In addition, the software developer understands the tests or verification events that must be passed for user acceptance.

- By integrating the MUIRS analyses into the ATAM scenario development, sustainability and safety/security aspects cannot easily be omitted from the system design. As the testing plan flows from the scenarios, the MUIRS design elements will have corresponding test or verification events identified in the test plan.

**Failure Modes and Effects Criticality Analysis**

As the title indicates, this analysis methodology is designed to identify system failure modes and those failures’ effects on the system, and ascertain the relative criticality of that type of failure. Blanchard (2004) described FMECA as follows:

> Given a description, both in *functional* and *physical* terms, the designer needs to be able to evaluate a system relative to possible failures, the anticipated modes and expected frequency of failure, their causes, their consequences and impact(s) on the system overall, and areas where preventative measures can be initiated to preclude such failures in the future. (p. 275)

He went on to state, “The FMECA is an excellent design tool, and it can be applied in the development or assessment of any product or process” (Blanchard, 2004, p. 276).

Including FMECA scenarios with the software systems and subsystems provides architectural design cues to software engineers. These scenarios provide analysis for designing redundant systems for mission-critical elements, provide “safe mode” operations for survivability- and safety-related systems, and drive the software engineer to conduct “what if” analyses with a superior understanding of failure-mode scenarios.

**Primary Software Acquisition Problem Areas Addressed**

The primary problem areas addressed by FMECA include requirements clarification and prioritization, and helping to ensure a sound software architecture design. This analysis also ensures that the most critical software systems are designed with the requisite reliability and will continue to function in degraded modes.

As previously stated, one of the main functions of performing FMECA is to identify those software functions that are not critical, and to ensure that failures or anomalies in those non-critical functions do not preclude or negatively affect system capabilities. Today’s systems typically have numerous enhancing functions that improve performance but are not critical, and the software developers have no way to discern the difference between a critical system and an enhancing one without employing FMECA.

**Integrating the Recommended Tools, Techniques, and Processes into the Defense Acquisition System**

The tools, techniques, and processes were specifically selected for both their ability to address software-intensive systems development problems and their ability to integrate with the DAS. They are all SEP enablers designed to improve the critical DAS front-end processes, which are primarily the government’s responsibility.

Figure 2 shows the processes applied at the latest possible developmental time to be effective. The earlier these tools, techniques, and processes occur, the more effective they
As depicted in Figure 2, the QAW and ATAM are designed to address critical requirements and design front-end processes, where the government is primarily responsible for the process. The blue arrow shows how the warfighters and user community are continuously involved throughout the process, and are active participants in the QAW and ATAM processes. This is distinctly different than the traditional DAS where there is little formal user interaction between preparation of the JCIDS documents and the prototype limited user tests (LUT)/early user test and evaluation (EUT&E).

The user communities have a very significant role in driving the QAW and ATAM processes, which requires more user resources to support the system development. This user investment in the DAS is becoming more critical with the development of more software-intensive and complex systems of all kinds. This investment is absolutely necessary to avoid government to contractor misunderstanding of the system requirements and warfighter expectations, and would significantly reduce the cost and schedule delays associated with user dissatisfaction, user-test failure, and unnecessary system redesign.

**Program Management Risk Reduction**

These tools, techniques, and processes will not, of themselves, produce or guarantee anything. “An architecture analysis method, any architecture analysis method, is a garbage-in-garbage-out process. The ATAM is no different. It crucially relies on the active and willing participation of the stakeholders” (Kazman et al., 2000, p. 63). All of the tools and techniques described and recommended in this research are dependent on the team of professional stakeholders conscientiously performing their critical function in the development of the software-intensive system.
To effectively implement the recommended tools, techniques, and processes, the program management team must be professional, disciplined in their application of the SEP, and skilled in integrating the tools into the DAS. In a word used by the SEI, the team must be *mature*. The Defense Acquisition Workforce Improvement Act (DAWIA) mandates certain education and training levels for the professional workforce performing at various levels. The DoD invests significant resources in both education and training to help ensure the acquisition workforce competencies and comply with the DAWIA.

The DoD also evaluates the maturity of potential software developers by requiring an evaluation using SEI’s Capability Maturity Model–Integrated (CMMI; or equivalent) for most software-intensive system acquisitions. The CMMI is a five-level model, and the software developer organization under evaluation must achieve at least a level three by an independent evaluation team to be eligible to be awarded the DoD contract.

As mentioned previously, the DoD does not currently require the PM offices managing software-intensive systems to achieve any maturity level on the Software Acquisition Capability Maturity Model (SA-CMM). The team effort between the government and the software developer strongly suggests that both the PM office and the software developer would reduce developmental risk by demonstrating an appropriate level of maturity.

Due in large part to the immature software engineering environment, each major DoD software design and build tends to be unique. That means that the software development in complex systems will act the same way as integrating a new technology would, and the resulting program risk is very high. The software TRLs have little meaning in this type of environment, so risk management is highly dependent on the government and software development teams’ abilities to manage the system software development as a new technology with a low TRL.

A significant portion of the risk management is focused on the government and software development teams. As the software TRLs are mostly ineffective, I would recommend the further development of software Management Readiness Levels (MgtRLs) to mitigate the risks. Part of the management risk reduction is already in place with the DAWIA requirements and the software developer maturity levels that must be achieved.

**References**


Investing in Software Sustainment

Robert Ferguson—is a senior member of the technical staff at the Software Engineering Institute (SEI). He works primarily on software measurement and estimation. He spent 30 years in the industry as a software developer and project manager before coming to the SEI. His experience includes applications in real-time flight controls, manufacturing control systems, large databases, and systems integration projects. He has also frequently led process improvement teams. Ferguson is a senior member of IEEE and has a Project Management Professional (PMP) certification from the Project Management Institute (PMI). [rf@sei.cmu.edu]

Abstract

In many government weapon systems, sustaining software depends heavily on organic engineering efforts. This is different from hardware sustainment (the more traditional form of sustainment), which often depends heavily on the supply chain and service providers and much less on engineering capability. Because of this shift, a larger portion of sustainment funding needs to be allocated to improving the sustainment infrastructure within government sustainment organizations. This includes the engineering processes, tools, and skills of engineering staff. Failure to recognize this need in a timely fashion has the potential to increase sustainment costs and, at the same time, degrade system performance.

The decisions and processes are complex because various stakeholders make decisions at different times, yet these decisions are interrelated, impact one another, and create constraints on the ability of the sustainment organization to fulfill its mission. To deal with the complexity of the decision-making process, the Software Engineering Institute (SEI) developed a simulation model for analyzing the effects of changes in demand for software sustainment and the corresponding funding decisions. The model allows decision-makers to analyze multiple allocation strategies in response to changes from mission command and budget authorities. The model has been tested and calibrated using historical data and is now in operational use by the Process Resource Team at the Naval Air Weapons Station China Lake.

Introduction

The ideas for this model were developed in response to real-world events. One DoD program charged with sustaining a 20+ year old aircraft system asked the SEI how it could justify the capital investment necessary to update its test and support systems and the supply chain. Some of the parts were past end-of-life, making them difficult to source, and the radar—a major technology component—had been updated without updating the test equipment. The situation was placing a major strain on the organization. Engineers had to spend significant work time on eBay buying parts. Radar testing costs escalated significantly because of the new steps required:

---

1 Copyright 2015 Carnegie Mellon University. No warranty. This Carnegie Mellon University and Software Engineering Institute material is furnished on an “as-is” basis. Carnegie Mellon University makes no warranties of any kind, either expressed or implied, as to any matter including, but not limited to, warranty of fitness for purpose or merchantability, exclusivity, or results obtained from use of the material. Carnegie Mellon University does not make any warranty of any kind with respect to freedom from patent, trademark, or copyright infringement. This material has been approved for public release and unlimited distribution. DM-0002313.
1. Remove an airplane from operations.
2. Disassemble the radar to put it in the lab.
3. Calibrate the radar in the lab setting.
4. Complete testing.
5. Reassemble the aircraft with the radar.
6. Recalibrate the radar to the airplane.

Repercussions of this problem will continue even after the equipment becomes available and a new source of parts is established. Both physical plant and process changes will be required to reestablish the level of productivity the organization had when its infrastructure matched the technology requirements of the aircraft systems being sustained.

This type of situation occurs with some frequency and can be summarized as follows:

- An organic sustainment organization with valuable facilities and skills is already in place.
- A new technology makes parts obsolete and leads to new engineering design work.
- The sustainment organization needs to upgrade both skills and facilities to meet the demand.
- While waiting for these changes to be completed, the sustainer’s efficiency is compromised, and “mission capable availability” is diminished.
- Until process and tools stabilize, quality often suffers.

Part of the problem for the sustainers is the familiar “color of money” problem; by law, specific funding sources must be applied to specific uses. Funding for product modernization is supplied by the acquisition budget and arrives via the program office. Funding for developing organic sustainment capability typically comes from the life-cycle command function. Delays in funding to update facilities and processes will eventually cause problems in mission performance. Detecting this situation is nearly impossible using spreadsheet analysis alone.

At least five distinct stakeholders are present in sustainment work:

1. Actual operators, who represent the mission-use viewpoint
2. Strategic planners, who review threats to the existing system and opportunities for new system capability based on changes in technology
3. Sustaining engineers, who must address requests for new capabilities as well as addressing the effects of external changes to existing subsystems (e.g., software changes to sensors or communications.)
4. The sustainers’ management team, which must invest in facilities and organizational capabilities and retain talent with product domain knowledge
5. The program office, which is responsible for the flow of funds and for promoting the program to all stakeholders

All of the stakeholders have their own definitions of value or utility. Each also has a different timeframe for decisions; hence, each stakeholder may perceive any of the others as delaying the response. The dynamic is one of constant change since technology may change and new threats arise at any time. Figure 1 shows the interaction of the stakeholders and the potential for various gaps in the desired performance.
Figure 1. Stakeholders and Gaps

Rationale for a Simulation

Sustainment is part of a dynamic and changing system. Different forces are stronger at different times, and there are time delays between decisions and outcomes that hide valuable information from decision-makers. Any thought of perfect decisions point toward an unattainable ideal. The purpose of the simulation model is to increase the likelihood of better decisions by forecasting several potential futures and examining the consequences. Even if the decision is not optimal at the start, the simulation indicates ways to make course corrections in a more timely fashion.

Simulations use scenarios of input and change to make these forecasts. Each scenario assumes a possible future budget allocation and a future set of demands on both the mission and the sustainer. The simulation then provides a graphical picture of future performance and potential gaps in that performance. In many cases, the simulation will show how a decision that appears sensible at first actually generates a significant performance gap within just a few months. Since some decision cycles are often two years or more, the gap may be one that cannot be overcome in a shorter timeframe than the next decision cycle. This information may be useful in persuading stakeholders (funders) to make different decisions about funding.

The Simulation Tool and Model

The model uses the systems dynamics method described in Peter Senge’s (2006) book *The Fifth Discipline*. The SEI’s simulation was developed using Vensim and systems dynamics methods (Ventana Systems, n.d.). This type of simulation uses a “stock and flow” model common to systems dynamics (Figure 2).
An epidemic is a good example because the nonlinear behavior and external interactions can create a future dangerous situation. The boxes indicate “stocks” that can be supplied or drained by “flows,” represented by double arrows. Stocks may also be initialized. The double triangle is a valve that controls the rate of the flow. The source of a single arrow is a measure. The target is a calculation or control of a valve. If the source of the arrow shows an increasing/decreasing value, the sign +/- implies that the target of the arrow has the same/opposite impact as the measure of the source. In the diagram, an increase in the susceptible population creates an increasing number of contacts. Similarly, if the duration of the infection grows, the recovery rate decreases (the “-”).

The full simulation for sustainment is quite large and can be seen in the SEI publication, *A Dynamic Model of Sustainment Investment* (Sheard et al., 2015). A smaller piece showing how technology demands interact with mission performance appears in Figure 3.
Figure 3. Mission and Strategy Portion of the Simulation Model

The Vensim user interface (presented in Figure 4) provides a customizable control panel with charts and slider bars. The charts show the performance graph of individual variables. The slider bars are a simple control panel for testing different inputs or input equations that drive the simulation. The charts can be selected from any number of available variables in the model. Additional scenarios can be developed to show how actions taken later affect results.

This interface allows decision-makers to use slider bars to set different funding allocation strategies and changes in demand and immediately see the effects of different actions.

Legend and Interpretation

In Vensim the +/- signs are replaced by “S”=+ and “O”=-.

The red colored arrows simply represent a reinforcing loop called the “Bandwagon Effect”. This effect simply means that a series of successful missions will increase the demand for future deployments and an expansion of mission goals. The demand then generates requests for new capabilities. In turn the additional capabilities result in greater expectations of mission performance.

Figure 4. Vensim User Control Interface

Legend and Interpretation

In Vensim the +/- signs are replaced by “S”=+ and “O”=-.

The red colored arrows simply represent a reinforcing loop called the “Bandwagon Effect”. This effect simply means that a series of successful missions will increase the demand for future deployments and an expansion of mission goals. The demand then generates requests for new capabilities. In turn the additional capabilities result in greater expectations of mission performance.
Calibration

Calibrating a simulation is always challenging. A model must be simpler than a real-world situation since modeling complete reality is far too expensive and time consuming. Each simplification involves some abstraction from reality, resulting in some redefinition of data. Sometimes an approximation or proxy must be used if real data is not available.

The most common way to calibrate the model is to begin by establishing an equilibrium: When the equilibrium is established, each stock and flow appears constant. This approach works because equilibrium can usually be established by manipulating a smaller subset of the variables and formulas.

A total of 26 flows and 19 stocks are defined in the SEI model. Actual calibration was performed with approximately half the total number of flows and just five of the stocks. The equilibrium values have to be reconciled with data observations of the real system. Two particular assumptions show the complexity of the abstraction:

- **Enhancement Requests:** Every enhancement request was counted as having the same size and effect on the developers. This is clearly inadequate for the longer term viability of the model. Requests come in different sizes—some are big and complicated, and some are much smaller and easier. For improved accuracy, the requests are considered as a set of sizes from very small (VS) to very large (VL).

- **Staff Capabilities:** It was also necessary to connect process capability and organizational capacity. We chose the simple formula:
  \[ \text{Sustainment capacity} = A \times \text{(number of capabilities)} \times \text{(number of staff)}, \]
  where A is some numeric value that helps to achieve equilibrium when calibrating. This formulation suggests that staff capabilities are closely related to process capabilities, which has been observed in many studies. Proceeding this way, we determined that a capability change could be based on training days, since process changes had to be supported by training days. The stable solution at this time is about 45 days per staff, or about 9,000 total capacity across an organization of 200 people. Capacity is diminished by staff members leaving or reductions in total staff. Improving capacity requires both staff and training employed together.

It is possible that neither of these assumptions will be valid after further study, but the simulation behavior appears acceptable in current use.

Developing Scenarios

Typically, the simulation should imitate real performance and should break where the system would break. Testing the simulation must include running scenarios of change and correlating the changed behavior of the model to the organization’s behavior. A scenario can be described using the elements in Table 1.
Table 1. Scenario Example

<table>
<thead>
<tr>
<th>ELEMENT</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario name and short description</td>
<td>New mission threat</td>
</tr>
<tr>
<td>Context of operation</td>
<td>Normal sustainment activity with a steady rate of enhancement requests</td>
</tr>
<tr>
<td>Stimulus</td>
<td>Aircraft squadron runs a normal mission to destroy anti-aircraft capability. However, opponent develops new (technology) capability to detect aircraft.</td>
</tr>
<tr>
<td>Response (or decision)</td>
<td>Mission planners request new technology for aircraft mission package to reduce ability to detect airplane. This additional work places a strain on the sustainment organization’s ability to deliver on time. Overtime is requested at 20% rate to cover additional demands.</td>
</tr>
<tr>
<td>Outcome (what the simulation tells us)</td>
<td>Productivity declines over time anyway.</td>
</tr>
</tbody>
</table>

Additional scenarios can be tested simply by changing any of the first three elements. For example, adding test equipment represents a potential response. Operating during wartime is a new environment. Finally, an example of a new stimulus would be an enemy’s new capability to detect a stealth aircraft. Similar scenarios may be generated by any input, whether from customer demand, operational performance, or sustainment performance. Each new scenario may have different effects on the various outcomes in terms of productivity, customer satisfaction, and the ability to perform missions.

Results

Four distinct scenarios were tested during the research. The results of the “Sequestration” scenario are described in this section.

In sequestration, the stimulus was a 20% cut in funding. The initial response was to cut 100% of the training budget. The outcome for that particular response was a steady decline in productivity for as long as sequestration lasted. An alternate response was also considered: maintain training and increase investment in process improvement. The alternate response demonstrated a short-term, six-month loss in productivity but showed higher levels of productivity within 12–18 months. Obviously, there are practical limits to improving productivity, but the model assumes that the capital investment is wisely spent.

Figure 4 also illustrates the possibility of considering alternatives based on changing assumptions about either the baseline or the stimulus event (represented by the box labeled “Alternate Baselines”).

The other three scenarios we investigated, which are described in the larger paper, were

- Gating the demand: The sustainment organization chooses to deny a number of requests because it does not have the capacity to fulfil them all. What will it do to sustainment performance and to the mission performance assessment?
- New threat, no budget: An opponent develops a technology that could compromise mission performance. The system is upgraded, but there may not be money for training and tooling. How long does it take for mission performance to decline and by how much?
• Underfunding sustainment investment: Underfunding occurs when the sustainers are made responsible for the support of a new technology, but there is no budgetary provision to support tooling, process changes, and staff training. Will this affect mission readiness, and how long will it be before the results are observed?

These initial scenarios are not completely independent. “New Threat” proved to be very similar to “Underfunding.” The advantage of documenting the scenario carefully is the ease with which different scenarios can be tested.

Outcomes and Future Work

The Process Resource Team and the Naval Air Weapons Lab worked with the SEI on this model. Their confidence in the model was sufficient enough for them to purchase Vensim and to seek a longer term relationship with the SEI to study and improve the model.

Several kinds of improvement to model are possible. The most obvious is to improve calibration, which will require some additional data. In particular, concepts associated with mission capability, mission performance, and demands for new capability need more precise definitions and better measurement data. Another abstraction that should be reconsidered would address potential trades between process capability and individual capability. The initial model assumes these are the same measure.

The Naval Air Weapons Lab plans to extend the use of the model to other programs within NAVAIR. The future practicality of the simulation model depends on both our ability to calibrate the model to the system of interest and whether the scenarios tested represent actual experience. In any case, the simulation will almost certainly help the organization develop better long-range plans and improve cost and schedule risk mitigation.

References


Acknowledgements

Erin Harper, Technical Editor, Software Engineering Institute, and Jeff Schwalb, Process Resource Manager, Naval Air Weapons Station China Lake.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.
Panel 13. DoD’s Contracting Portfolio: Analysis and Innovations

Thursday, May 14, 2015

11:15 a.m. – 12:45 p.m.

Chair: Seán F. Crean, RADM, SC USN (Ret.), Director, Office of Government Contracting at U.S. Small Business Administration

- Andrew Hunter, CSIS
- Gregory Sanders, CSIS
- Jesse Ellman, CSIS

**Contractor Past Performance Information: An Analysis of Assessment Narratives and Objective Ratings**
- Rene Rendon, NPS
- Uday Apte, NPS
- Michael Dixon, NPS

**An Open Door and A Leg Up: Simplified Contracts for Service-Disabled Veteran-Owned Small Businesses at the Department of the Navy**
- Max Kidalov, NPS

Seán F. Crean—is the director of the Office of Small Business Programs for the Department of the Navy. He serves as chief advisor to the secretary on all small business matters. He is responsible for small business acquisition policy and strategic initiatives.

Crean joined the Secretary of the Navy staff as a member of the Senior Executive Service in January 2010 and has over 30 years of federal service. Prior to receiving this appointment, he served as deputy assistant secretary of the Navy for Acquisition and Logistics Management during a two-year military recall to active duty as a rear admiral in support of Operation Iraqi Freedom.

Crean’s previous experience includes serving as the senior procurement analyst for the U.S. Small Business Administration’s Office of Government Contracting Area I (New England) for 19 years. In this role, he was the principal advisor to the SBA’s six regional district offices and congressional delegations on procurement issues. He provided acquisition strategy analysis for over 20 buying activities throughout the region, supporting both DoD and civilian federal agencies. He first entered federal civilian service as the deputy supply officer for Naval Air Station Brunswick, ME, where he was also appointed the activity small business specialist.

Crean’s combined military and civil service careers have provided complementary and extensive leadership responsibilities in service to the country. As a member of the reserve component, he has attained the grade of rear admiral (two-star) and is currently assigned as deputy commander, Naval Supply Systems Command. He holds a Bachelor of Science degree in business management and marine transportation from State University of New York Maritime College and a Master of Business Administration degree from New Hampshire College’s Graduate School of Business.

He has a number of personal and command decorations, including two Legion of Merit awards. He is a member of the Defense Acquisition Corps and is DAWIA Level III Contracting certified.

Andrew Hunter—is a senior fellow in the International Security Program and director of the Defense-Industrial Initiatives Group at CSIS. From 2011 to 2014, he served as a senior executive in the Department of Defense, serving first as chief of staff to under secretaries of defense (AT&L) Ashton B. Carter and Frank Kendall before directing the Joint Rapid Acquisition Cell. From 2005 to 2011, Hunter served as a professional staff member of the House Armed Services Committee. Hunter holds an MA in applied economics from the Johns Hopkins University and a BA in social studies from Harvard University. [ahunter@csis.org]

Gregory Sanders—is a fellow with the Defense-Industrial Initiatives Group at CSIS, where he manages a team that analyzes U.S. defense acquisition issues. Utilizing data visualization and other methods, his research focuses on extrapolating trends within government contracting. This requires innovative management of millions of unique data from a variety of databases, most notably the Federal Procurement Database System, and extensive cross-referencing of multiple budget data sources. Sanders holds an MA in international studies from the University of Denver and a BA in government and politics, as well as a BS in computer science, from the University of Maryland. [gsanders@csis.org]

Jesse Ellman—is a research associate with the Defense-Industrial Initiatives Group (DIIG) at CSIS. He specializes in U.S. defense acquisition policy, with a particular focus on Department of Defense, Department of Homeland Security, and government-wide services contracting trends; sourcing policy and cost estimation methodologies; and recent U.S. Army modernization efforts. Ellman holds a BA in political science from Stony Brook University and an MA with honors in security studies, with a concentration in military operations, from Georgetown University. [jellman@csis.org]

Abstract

Existing research on defense products contracting shows a marked focus on large and high-profile programs that qualify as Major Defense Acquisition Programs (MDAPs). In an effort to fill a gap in the literature of product contracting, CSIS has launched this study—an analysis of trends in DoD products contracting covering all contracting types, key components accounts, and categories of products.

Looking at the period from 1990–2014, this report presents initial results and focuses on three notable questions: How have rates of effective competition differed between different categories of products throughout the period? How have the industrial bases for different categories of products changed since 1990? And how did sequestration and its aftermath affect DoD products contracting? Additionally, this report identifies notable data quality issues with both the Federal Procurement Data System and DD350 data to aid future research.

The main findings of this initial inquiry are threefold: first, that “Last Supper” industry consolidation has affected the vendor size mix and levels of competition in the defense products industrial base to this day; second, that sequestration has had profound effects on what products the DoD buys; and third, that contracting trends vary significantly for the various DoD components and categories of products.
Introduction

For almost a decade, the Defense-Industrial Initiatives Group at the Center for Strategic and International Studies (CSIS) has analyzed and reported on trends in federal contracting in general, and Department of Defense (DoD) contracting specifically. While initially focused on services contracting as well as the overall trends of DoD contracting, CSIS has in recent years sought to drill down deeper into this data with regard to the DoD’s key components, and for the different varieties of goods and services that the DoD contracts for. Past CSIS work for the Naval Postgraduate School has focused on trends in DoD services contracting; in this report, the study team shifts its focus to DoD products contracting.

The existing literature on defense products contracting tends to focus on large, high-profile and high-cost programs that qualify as Major Defense Acquisition Programs (MDAPs). This focus, however, does not capture the full range of products that the DoD contracts for. In an effort to fill a gap in the literature, CSIS has undertaken a comprehensive analysis of trends in DoD products contracting for the DoD and its key components: the Army, Navy, Air Force, Defense Logistics Agency (DLA), and “Other DoD,” which comprises all remaining contracting entities not captured by the first four categories.

In order to facilitate analysis of the differences in contracting trends between different types of products, CSIS has created a taxonomy of the universe of DoD products, using U.S. government Product and Service Codes (PSCs) to separate DoD products into 10 product categories: Aircraft, Clothing & Subsistence, Electronics & Communications, Engines & Power Plants, Fuels, Ground Vehicles, Launchers & Munitions, Missiles & Space, Ships, and “Other.”

This report presents interim findings from this analysis. Though the overall research effort will delve deeply into the product categories that are not typically associated with MDAPs, this initial paper will focus primarily (but not exclusively) on the more prominent product categories, as the study team attempts to validate the overall research methodology and better understand the particular issues involved with the less prominent categories of defense products.

This paper focuses on three areas related to both historical and recent trends in defense products:

---

1 The Center for Strategic and International Studies (CSIS) does not take specific policy positions; accordingly, all views expressed in this presentation should be understood to be solely those of the author(s).
2 Though the CSIS DoD products taxonomy does not necessarily align with the taxonomy developed within the DoD, CSIS does have the capability to cross-walk between the two.
3 The five product categories that are not mostly comprised of platforms and programs related to MDAPs (Clothing & Subsistence, Electronics & Communications, Fuels, Launchers & Munitions, and “Other”) accounted for 42% of DoD products contracts in 2014, though those categories do include some MDAPs. Using the FPDS system equipment code field, over 60% of contract obligations are not associated with an MDAP in 2014, but that field is not filled in consistently. Though neither of these methods can provide a precise figure for the share of products contract obligations associated with MDAPs, it is safe to say that a significant share of defense products contract obligations, perhaps as much as half, are not associated with MDAPs.
Historical trends in competition for DoD products
Changes in the industrial base for DoD products
The impact of sequestration and its aftermath on DoD products contracting overall, by component, and by product category

Methodology

To provide greater historical context to recent trends, CSIS has integrated fiscal year (FY)1990–1999 contracting data into its analysis for this study. All data from FY2000–2014 is drawn from the publicly-available Federal Procurement Data System (FPDS) through the USASpending.gov portal. Due to a lack of pre-2000 data available through USASpending.gov and how unwieldy it is to get the full range of relevant study variables for the entire department using the FPDS.gov web tool, CSIS is using archival DD350 data for the 1990–1999 period. The adoption of archival DD350 data for the period 1990–1999 poses challenges of which we are aware and have worked diligently to mitigate and standardize.

This report relies on the methodology that the study team has established and refined for analysis of federal contracting data over the course of the last decade. For this study in particular, there are a few key differences and updates:

- All dollar figures are in constant 2014 dollars, using the latest OMB deflators.
- In FY2013, the Defense Commissary Agency (DeCA) stopped reporting most of its contract obligations (approximately $5 billion) into FPDS. Because this creates a significant data discrepancy that distorts trend analysis, CSIS has excluded DeCA from the data set throughout the period.
- For analysis of the industrial base, the composition of the “Big 6” defense vendors has changed—BAE Systems, which has declined as a DoD vendor in recent years, has been replaced by United Technologies for all years in the data set.

Notable Limitations and Gaps in Pre-2000 Contract Data

Use of archival DD350 data for the 1990–1999 period carries some cost in data quality, as there are notable differences in coding schema and granularity between the DD350s and the modern FPDS architecture. The following are the most notable issues:

- DD350 data for FY1990–FY1999 reflect pre-FY2004 reporting thresholds, which did not require the DoD to report more than summary information on contracts below $25,000.
- FY1990 has a significant percentage of data left blank or otherwise unclassifiable, mostly in the fields used for competition, pricing mechanism, and vehicle.

---

4 Past CSIS work has at times included 1990–1999 data extracted from the FPDS.gov web tool, but that approach did not allow for examining vendor size or examining more than one variable at a time.
5 See http://csis.org/program/methodology for the complete methodology.
FY1994 data had a serious data issue wherein nearly all Army contracts were improperly classified under other components. CSIS has been able to partially correct this issue, and is continuing to seek a full solution, but Army contract obligations for 1994 remain understated.

The DD350 does not include the “Statutory Exemption to Fair Opportunity” field, which CSIS uses for greater precision on levels of competition for Indefinite Delivery Vehicle (IDV) contracts.

Prior to FY1997, DD350 data did not reliably differentiate between numbers of offers greater than two (such that most contracts receiving two or more offers had “2” listed under number of offers). As such, pre-1997 competition data has reduced granularity in terms of number of offers.

Attempts to use data from FPDS.gov to address these issues have been hampered by a more serious data gap: For 1990–1994, the total DoD contract obligations in FPDS are approximately $20 billion per year lower than in the data contained in the DD350s, representing about a sixth of total DoD contract obligations for those years. Upon further investigation, the study team found that a number of large contracts in the DD350 data set are either completely missing from FPDS or have vastly lower obligation levels associated with them. CSIS is currently engaged with policy makers inside the DoD to raise awareness of this issue, identify the source of the data gap, and work toward a solution.

Though these are serious data quality issues, the CSIS nonetheless believes the overall quality and reliability of the data set is more than sufficient to perform meaningful trend analysis.

**Historical Trends**

**Competition for Defense Products**

In a recent short paper, the study team noted that competition rates for products, services, and R&D for the DoD overall were remarkably consistent from 2008–2014 (Ellman, 2014). During that period, around one-third of contract obligations were awarded after effective competition. With the integration of pre-2000 data into the CSIS federal contracting database, the study team can now extend this analysis back to 1991. This historical data gives us new insight into the Post–Cold War drawdown, a period that gives important context to the present drawdown. This comparison is particularly valuable because it can help illuminate the effects of sequestration, which was not a factor in the 1990s.

From 1991 to 1998, the rate of effective competition for DoD products contracts fluctuated between 38% and 42% in all but one year (36% in 1995). The data shows a slow but steady decline in effective competition rates after 1996, which coincides with the Last Supper industry consolidation that removed a number of major competitors from the defense market via mergers and acquisitions.

---

6 CSIS defines *effective competition* as a competitively solicited award that received two or more offers, which is similar to the DoD’s definition of effective competition.

7 Data from 1990 are excluded here because nearly 36% of obligations in that year are “unlabeled” for competition.
The rate of effective competition fell from 40% in 1996 to 33% in 2004, before surging to 40% in 2005. From 2007 on, effective competition rates hovered between 33% and 37%.

The rate of effective competition for overall DoD products does not tell the whole story because there are significant differences between the different product categories. Figure 1 puts these differences into context by comparing 2014 competition rates for each product category to their overall average competition rate from 1991–2014.

![Figure 1. 1991–2014 Effective Competition Rates vs. 2014 Effective Competition Rates for Defense Products by Product Category](image)

Note. Data in this figure comes from the FPDS, CSIS analysis.

**Figure 1. 1991–2014 Effective Competition Rates vs. 2014 Effective Competition Rates for Defense Products by Product Category**

For overall DoD products, the 2014 effective competition rate is right in line with the rate of effective competition for the entire period, reinforcing the point that the rate of competition for products has been overall quite stable over time. For Aircraft, Missiles & Space (M&S), Ground Vehicles, Electronics & Communications (E&C), and Fuels, the 2014 rates of competition are within a few percentage points of their historical averages. Rates of effective competition in 2014 for Engines & Power Plants (E&PP), Launchers & Munitions (L&M), Clothing & Subsistence (C&S), and “Other” are notably above their historical averages, while the effective competition rate for Ships is well below the historical average.

However, historical averages can only provide so much visibility into trends, and effective competition rates within product categories have been quite volatile from year to year; given the relatively “small” size of the various products categories, it is unsurprising that one or two big contracts (or a group of smaller ones) can cause big shifts in effective competition rates. Still, looking at effective competition rates over the period in question can provide greater insight into long-term trends.

**Platforms and Complex Systems: Aircraft, Ground Vehicles, Ships, and Missiles & Space**

For this analysis, the study team has focused on trends in the rates of effective competition for the four platform-focused product categories: Aircraft, Ground Vehicles,
Ships, and Missiles & Space. Figure 2 shows effective competition rates for those four product categories between 1991 and 2014.8

Figure 2. Rate of Effective Competition for Aircraft, Ground Vehicles, Ships, and Missiles & Space, 1991–2014

**Aircraft**

The effective competition rate for Aircraft has consistently been among the lowest, if not the lowest, among the product categories, owing both to the length and complexity of the programs and the limited number of vendors able to effectively compete at the development stage. Interestingly, even after the Last Supper–inspired flurry of mergers and acquisitions in the mid-1990s radically changed the landscape of the Aircraft industrial base, the rate of effective competition was largely stable in the low-to-mid 20s until 2003, when the rate fell to 14%. One possible explanation for the delayed effect is the amount of time it takes for new major Aircraft programs to start up: Large programs like the F-22 and F-35 were just starting to ramp up in the early 2000s. After 2003, the rate of effective competition for Aircraft has never exceeded 13%, hovering between 10 and 12% in most years, about one-third the rate for overall DoD products.

**Missiles & Space**

Similar to Aircraft, the consistently low rates of effective competition for M&S are reflective of highly complex programs with a limited industrial base. The rate of effective competition for M&S peaked at 36% in 1992, but dropped off steadily (with a couple of plateaus) afterwards, to a low of 11% in 2000. The rate spiked up briefly to between 19%

---

8 Data for 1990 is excluded due to data quality issues. See the Methodology section for further discussion.
and 22% from 2002–2005, primarily due to increased Air Force contract obligations for “space vehicles,” but has dropped off since, and did not exceed 12% in any year from 2007–2013. In 2014, the rate of effective competition doubled, from 7% in 2013 (the lowest rate in the period) to 14%.

**Ground Vehicles**

The available industrial base able to compete for most ground vehicle programs is significantly broader than that for the two previous categories, so it is not surprising that the rate of effective competition for ground vehicles has usually been somewhat higher. That rate was extremely volatile in the 1990s, rising as high as 44% in 1995, and then falling back to 17% in 1999 and 2000. That volatility continued into the early 2000s, with effective competition rates rising to 41% in 2001, and then fluctuating between the mid-to-high 30s and the high teens over the next four years. Between 2006 and 2012, however, effective competition rates have hovered in the low-to-mid 30% range, in part as a result of the highly competitive MRAP contracts in the mid-to-late 2000s. The effective competition rate for Ground Vehicles fell from 34% in 2012 to 25% in 2013, but rose back to 29% in 2014.

**Ships**

Considering the size and complexity of the platforms involved, the high rates of effective competition for Ships in the 1990s—as high as 91% in 1992, and over 60% in all but one year between 1991 and 2000—is somewhat surprising. The private shipyard consolidation in the late 1990s and early 2000s had a major impact on that rate, as the rate of effective competition never exceeded 50% after 2002, only exceeded 30% in one year after 2005, and remained below 20% from 2006 through 2009. Since 2009, however, the rate of effective competition has increased, remaining above 25% in every year and rising to 32% in 2013. This is likely the result of deliberate decisions to split the procurement of certain high-cost platforms, such as the Littoral Combat Ship (LCS) and Aegis-class destroyers, between two competing shipyards.

**Changes in the Defense Products Industrial Base, 1990–2014**

In addition to tracking trends on the customer side of defense contracts, CSIS also tracks trends on the vendor side of the equation. In particular, the study team has built up the capability to track contract obligations by size of vendor. In order to facilitate this analysis, vendors are divided into four size categories: Small, Medium, Large, and the Big 6. Any organization designated as small by the FPDS database—according to the criteria established by the federal government—was categorized as such unless the vendor was a known subsidiary of a larger entity. Vendors with annual revenue of more than $3 billion, including from nonfederal sources, are classified as large. The Big 6, broken out from “large,” consists of the six largest defense firms in recent years (Boeing, Lockheed Martin, Northrop Grumman, General Dynamics, Raytheon, and United Technologies Corporation). And any contractor that qualifies as neither small nor large is classified as “medium.”

---

9 United Technologies replaces BAE in the Big 6 in this report and going forward because BAE has fallen off in recent years as DoD purchases of ground vehicles have slowed.

10 While Northrop Grumman, United Technologies, Raytheon, and General Dynamics were all consistently among the top defense products vendors before the Last Supper industry consolidation,
The shares of DoD products contract obligations awarded to each size category (plus the share marked as “Unlabeled” due to data quality issues) are shown in Figure 3.

**Figure 3. Share of Defense Products Contract Obligations by Size of Vendor**

The share of overall defense products contract obligations awarded to small vendors has been extremely stable throughout the period observed. Small vendors accounted for between 10% and 11% of products contract obligations in every year from 1991–2001, rose steadily to 15% by 2005, and gradually fell back to 11% from 2011–2013. The 14% share awarded to small vendors in 2014 is the highest since 2005.

The share of products contract obligations awarded to medium vendors hovered in the mid-to-high 20s in the early 1990s, peaking at 30% in 1995. That rate sharply declined in 1996, to 23%, and declined steadily for the rest of the decade; this is likely the result of the mergers and acquisitions coming out of the Last Supper. The share awarded to medium vendors never exceeded 19% after 2001, and has hovered in the mid-teens for most of the 2000s and 2010s, falling to an all-time low of 14% in 2014.

Boeing was not consistently in the top echelon of vendors until the merger with McDonnell Douglas (which had consistently been a top products vendor), and “Lockheed Martin” did not exist until the merger of Lockheed and Martin Marietta. For this analysis, both Lockheed and Martin Marietta are classified as “Large” in the years before they merged, rather than as Big 6, and Boeing is considered part of the Big 6 despite being a smaller player in the defense market than McDonnell Douglas before the merger. CSIS will consult with experts in the coming months to get a better understanding of the pre-Last Supper defense industrial base in order to determine if there is a better way to track contract obligations going to the largest defense firms in those years.
The shares of products contract obligations awarded to large and Big 6 vendors showed the biggest changes from the Last Supper industry consolidation of the mid-to-late 1990s. The share awarded to the Big 6 vendors nearly tripled between 1993 and 1999, rising from 15% to 44%; the merged Lockheed Martin accounted for 11% of the overall 1999 DoD products market. Meanwhile, the share awarded to large vendors fell by more than half over the same period, from 44% in 1993 to 21% in 1999. The difference in the magnitude of the changes is explained by the fact that some medium vendors were also involved in the flurry of mergers and acquisitions, shifting some products obligations to the Big 6.

In the 2000s, that trend began to reverse somewhat, with obligations to large vendors rising from 22% in 2001 to 35% in 2008, while the Big 6 share of DoD products contract obligations fell from 44% to 35% over that same period. Since 2008, the trend has shifted again, with the share going to the Big 6 rising steadily (to 43% in 2014), while the share going to large vendors declined to 30% by 2014. This change happened despite the countervailing trend of divestments, including Northrop Grumman’s divestment of its shipbuilding business into Huntington Ingalls Industries.

As with the other areas of analysis in this report, looking at overall trends for vendor size does not tell the whole story—the industrial bases serving the different product categories are vastly different. The following sections will look at select product categories that have seen interesting trends within their respective industrial bases.

**Aircraft**

The effect of the Last Supper industry consolidation was particularly pronounced for Aircraft. In 1995, 68% of contract obligations for Aircraft went to large vendors, while only 21% went to the Big 6. By 1999, large vendors accounted for only 11% of Aircraft contract obligations, while the Big 6 accounted for 79%. Since 1999, over 70% of Aircraft contract obligations have been awarded to the Big 6 in all but two years (66% in 2010, and 69% in 2011).

**Clothing & Subsistence**

There has been a notable shift in the industrial base for C&S over the 1990–2014 period. In the early-to-mid 1990s, small and medium vendors dominated the C&S market, with small vendors capturing over 40% of C&S contract obligations in most years, and medium vendors accounting for as much as 56%. After 1995, however, large vendors captured increasing shares of C&S contract obligations, rising from 4% in 1995 to 33% in 2002, with the majority of the increase drawn away from medium vendors, which fell to 35% by 2002.

In the years since, the share of C&S obligations awarded to large vendors continued to rise to 50% in 2009, drawing mostly from small vendors, which declined as a share of C&S obligations from 32% in 2002 to 20% in 2009, and have never exceeded 20% since.

**Electronics & Communications**

The Last Supper industry consolidation had a profound effect on the E&C industrial base, with the share of contract obligations going to the Big 6 rising from 11% in 1993 to 36% in 1999, drawing roughly equally from medium and large vendors, though large vendors recovered some of their share of the E&C market from the Big 6 in the early-to-mid 2000s. The most interesting trend in recent years has been the consistent growth in E&C contract obligations to small vendors: From 17% in 1999, the share awarded to small vendors has grown steadily through the intervening years, to a high of 29% in 2014, the highest share of any size category and over double the rate for overall defense products.
This growth has apparently come at the expense of the Big 6, which saw its share of E&C contract obligations decline from 36% in 1999 to 24% in 2014.

**Launchers & Munitions**

As with E&C, the industry consolidation in the wake of the Last Supper led to significant changes in the L&M industrial base. Between 1996 and 1997, L&M contract obligations to the Big 6 rose from 13% in 1996 to 45% in 2007, drawing roughly equally from medium and large vendors. In the mid-2000s, there was a notable surge in contract obligations to small vendors, whose share of L&M contract obligations rose from 9% in 2002 to 20% in 2007. That increase was short lived, however, and the share awarded to small businesses has declined steadily since, back to 9% in 2013 and 2014.

**Missiles and Space**

As with several other categories, M&S saw a post–Last Supper shift in contract obligations from medium and large vendors to the Big 6. Available data suggests that throughout the 1990–2014 period, the share of M&S contract obligations awarded to small vendors never exceeded 3%, by far the lowest share for small vendors of any product category. Similarly, while medium vendors accounted for shares in the mid-teens to low 20% range from 1990–1995, medium vendors have not been a significant factor in M&S contracting since 1999. Between 2000 and 2014, the share of M&S contract obligations awarded to medium vendors has only exceeded 4% in two years (6% in 2002, and 5% in 2014).

**The Impact of Sequestration on Defense Product Contracts**

This section will examine the impact of sequestration, and its aftermath, on DoD contract obligations for products overall, and for those DoD components and product categories that showed notable trends. In order to examine the sources of changes in obligations levels, the study team has done further analysis to examine the specific PSCs and system equipment codes (which identify the program a contract is associated with) which show notable changes in obligations levels from year to year.

**Trends in Overall Defense Contract Obligations Under Sequestration**

Even in the context of a sharp downturn in defense contract obligations since 2008, the decline in the last two years, as the DoD has had to live under sequestration and its aftermath, has been significant. Overall defense contract obligations have declined by 31% since 2008, from $409 billion to $283 billion, but nearly two-thirds of that decline (65%) took place in 2013 and 2014. Overall defense contract obligations declined by 15% in 2013, and fell a further 9% in 2014. The latter decline was particularly notable because of the perception that the decline in 2013 was heavily driven by work being delayed and pushed back into FY2014 in the midst of the uncertainty surrounding sequestration. This perception led many to believe that 2014 would see, if not an increase, then a stabilizing of overall defense contract obligations; instead, 2014 saw another sharp decline.

In the FY2013 edition of CSIS’ series of reports on DoD contract trends (Berteau et al., 2014), the study team noted that the dramatic decline of overall DoD contract obligations in 2013 was not evenly distributed among the major DoD components: The Army (-21%), Air Force (-22%), and DLA (-23%) all declined more rapidly than did overall DoD contracts,
while the Navy (-2%) and “Other DoD” (-8%) were relatively spared. Similarly, in 2014, the Army (-14%) declined notably more steeply than did the overall DoD, while “Other DoD” (-5%) declined more slowly than overall, Air Force contract obligations remained steady (0%), and the Navy (-11%) and DLA (-7%) declined at rates comparable to overall DoD.

In 2013, as overall defense contract obligations declined by 15%, defense products contract obligations declined roughly in parallel (-14%). In 2014, as overall defense contract obligations declined by 9%, defense products contract obligations again declined by 14%. Since their peak in 2008, defense products contract obligations have declined by 37%—the declines in 2013 and 2014 have accounted for 60% of that total decline.

**Defense Products Contract Obligations Within Major DoD Components (After Sequestration)**

As with overall DoD contracts, the declines in defense products contract obligations were not evenly distributed among the major DoD components, as seen in Figure 4.

**Army**

Army products contract obligations, which peaked in 2008, have declined by two-thirds since then.

---

11 The decline in “Other DoD” cited here is notably lower than was discussed in the aforementioned CSIS report due to the study team’s removal of DeCA data from the sample. This change was made in response to DeCA stopping reporting of most of its contract data into FPDS in 2013. See the Methodology and Data Issues section for further discussion.
In 2013, Army products contract obligations declined by 27%, twice the rate of overall DoD products. There were numerous drivers of this sharp decline, including a nearly $800 million decline in contract obligations related to the CH-47D helicopter, an over $1 billion decline related to the Tactical UAV program, a $1.7 billion decline in obligations for “rotary wing aircraft,” a $1.3 billion decline in contract obligations for “combat assault and tactical vehicles,” and a $500 million decline in contract obligations for “land mines.” At the same time, the Army did see significant increases in contract obligations for certain programs and types of products, including a $900 million increase in contract obligations related to the CH-47F helicopter, \(^{12}\) a $1.3 billion increase for the Longbow Apache Block III, largely for “airframe structural components,” and a nearly $600 million increase in contract obligations for “wheeled trucks and truck tractors.”

In 2014, Army products contract obligations declined by a further 15%, but this decline was roughly in line with the decline in overall DoD products. Significant declines were seen in all three cases noted above that saw significant increases in 2013: the CH-47F helicopter (-$700 million), Longbow Apache Block III (-$1.5 billion, to less than $20 million), and “wheeled trucks and truck tractors” (-$700 million). These declines were counterbalanced by significant increases in some programs and product types, including $700 million for the AH-64A Apache and $1.1 billion for the Scout helicopter program (from $42 million in 2013).

This analysis shows the significant degree of volatility in contracts for Army rotary aircraft programs, with obligations spiking and dropping off dramatically over the course of just two years.

**Navy**

Navy products contract obligations have been relatively preserved in the current budgetary downturn—as overall DoD products contract obligations declined by 37% between 2008 and 2014, Navy products only declined by 17%.

In 2013, as overall products contract obligations declined by 14%, Navy products contract obligations actually increased by 9%. The main driver of this increase was a nearly $8 billion increase related to the F-35 Joint Strike Fighter, but there were numerous other significant increases, including $2 billion related to the Patriot missile program, nearly $1 billion for both nuclear reactors and unspecified “combat ships and landing vessels,”\(^{13}\) an $800 million increase related to the CVN-68 aircraft carrier,\(^{14}\) $700 million related to the DDG-51 destroyer program, and $800 million for the H-1 helicopter upgrade program. Navy products also saw programs and types of products with significant declines in 2013,

---

\(^{12}\) The concurrent decline in obligations coded as related to the CH-47D and increase in obligations coded as related to the CH-47F may reflect a coding change, rather than a real change in contracting activity. The study team suspects that some CH-47F contract obligations were improperly coded as related to the CH-47D prior to 2013.

\(^{13}\) This report uses “unspecified” to refer to contracts that are not associated with a system equipment code in cases where the study team believes that a significant share of the obligations are related to an MDAP.

\(^{14}\) The study team believes these obligations labeled as related to the older Nimitz-class aircraft carriers, which are coded under the PSC for “Combat Ships & Landing Vessels,” may be mislabeled obligations tied to the newer Ford-class aircraft carrier program.
including a $3 billion decline related to the P-8 Poseidon aircraft/MMA, a decline of $2 billion related to the America-class amphibious assault ship, $1 billion for unspecified “combat assault and tactical vehicles,” and over $600 million for “miscellaneous vessels.”

In 2014, Navy products contract obligations saw a marked drop after the notable increase in 2013, declining by 17%, slightly more steeply than overall DoD products. Again, the main driver of the year-to-year change was in the F-35 program, which declined by over $8 billion to below the 2012 obligations level. Other programs and product types also saw significant drops in contract obligations, including a $700 million decline in the E-2C/E-2D Advanced Hawkeye program, a $500 million decline for the H-1 helicopter upgrade program, and a $1.4 billion decline in the LPD-17 amphibious transport dock program. Even with the steep decline, some Navy programs saw significant increases, including a $600 million increase related to the Ford-class aircraft carrier program and a $3 billion increase in the SSN-774 Virginia-class submarine program.

Overall, while there has been significant fluctuation in the contract obligations going to individual programs and product types, Navy obligations for products have been relatively steady once the volatility in the F-35 program is accounted for. For the most part, the Navy has been able to preserve funding for its key platforms despite the budgetary constraints imposed by sequestration in 2013 and its aftermath in 2014.

**Air Force**

Air Force products contract obligations declined somewhat more slowly than did overall DoD products during the current budget downturn—27% for the Air Force between 2008 and 2014, versus 37% for overall DoD products.

Air Force contract obligations were higher in 2012 than they had been since 2007, but declined by 28% between 2012 and 2013, double the rate of overall DoD products. The main driver of the decline was a $3.4 billion drop in obligations related to the C-17A transport aircraft program, as well as a $3 billion decline in obligations for unspecified “aircraft, fixed wing.” The most significant increase in Air Force contract obligations in 2013 was a $900 million increase in contract obligations related to the Shillelagh anti-tank missile, a 1970s Army program. CSIS believes this anomaly is due to the Air Force re-using system equipment codes, which are used in FPDS to tie contract obligations to a particular program. The money associated with the Shillelagh missile in FPDS is otherwise classified as “aircraft, fixed wing,” and CSIS is seeking an updated system-equipment codebook.

In 2014, Air Force products contract obligations declined by only 9%, versus 14% for overall DoD products. Three programs saw particularly significant increases in contract obligations in 2014: obligations related to the C130-J transport aircraft increased by $900 million, the JASSM cruise missile program saw a $450 million increase, and the NAVSTAR GPS satellite program saw a $350 million increase. Three Air Force programs or product

---

15 There was also a nearly $2.3 billion decline in Air Force products contract obligations related to the Evolved Expendable Launch Vehicle (EELV) program, but that decline was due to a coding change, with space launches related to the program being reclassified as a service, rather than a product.

16 Due to poor data labeling, FPDS shows almost no Air Force contract obligations associated with the F-35 program. The study team believes that the majority of unspecified “aircraft, fixed wing” obligations are related to the F-35 program.
types saw declines of approximately $400 million each: the AMRAAM missile program, “miscellaneous aircraft accessories and components,” and “electronic countermeasure and quick reaction equipment.” Notably, contract obligations for “airframe structural components” related to the A-10 Warthog close-air support aircraft fell by over $200 million to just $13 million in 2014, though this may be due to the timing of contracts, since a similar obligations level was seen in 2012 (after obligations levels near $200 million in 2010 and 2011).

Overall, the study team is wary of reading too much into the trends beneath the surface of Air Force products contract obligations in recent years, due to the significant data labeling issues. CSIS urges Air Force policy makers to promote greater clarity in the use of system equipment codes among those responsible for entering data into FPDS.

Defense Products Contract Obligations by Type of Product

As described in the methodology, CSIS has sorted the range of products for which DoD contracts into 10 categories in order to facilitate analysis of how contracting trends differ across differing types of products. Figure 5 shows how these 10 categories of DoD products have fared during the current budgetary downturn and in the wake of sequestration.

![Defense Products Contract Obligations by Product Category, 2008–2014](image)

**Note.** Data in this figure comes from the FPDS, CSIS analysis.

**Figure 5.** Defense Products Contract Obligations by Product Category, 2008–2014

---

In order to avoid repetition, the specific dollar changes to programs/specific product types mentioned in the “By Component” section above will not be repeated in this section, except where changes in cross-component totals differ significantly from the changes discussed above.
The following sections will look more deeply into trends for six of these product categories to determine the specific programs and product types that experienced significant changes in obligations levels under sequestration and its aftermath.

**Aircraft**

As overall DoD contract obligations declined by 37% since 2008, contract obligations for Aircraft have fallen at half that rate (-18%).

In 2013, despite the steep drop in overall DoD products under sequestration, contract obligations for Aircraft fell by only 2%, one-seventh the rate of overall DoD. As discussed in the component-specific discussions above, 2013 saw a near-tripling of obligations related to the F-35 Joint Strike Fighter, along with major increases related to the CH-47F helicopter, E-2C/E-2D Advanced Hawkeye, H-1 helicopter upgrade program, and Longbow Apache Block III. There were also sharp declines in obligations related to the C-17A transport aircraft, P-8 Poseidon aircraft/MMA, and Tactical UAV. Additionally, there was an approximately $800 million decline in obligations related to the V-22 Osprey across the services, a $2.9 billion decline in unspecified “aircraft, fixed wing” (primarily in the Air Force, and believed to be mostly unspecified F-35 contracts), and a $1.7 billion decline in unspecified “aircraft, rotary wing” (primarily in the Army).

In 2014, Aircraft contract obligations (-29%) declined at over twice the rate of overall DoD products, the largest decline of any product category in 2014. This decline was largely driven by the huge drop in F-35 contracts discussed in the Navy section above, which appears to be primarily an issue of contract timing. There were also significant declines in that same pool of Air Force unspecified “aircraft, fixed wing” contracts, as well as with contracts related to the CH-47F helicopter, E-2C/E-2D Advanced Hawkeye, H-1 helicopter upgrade, and Longbow Apache Block III, all of which saw significant increases in 2013. There were, however, some programs that saw significant increases in obligations in 2014, including the AH-64A Apache helicopter, C130-J transport aircraft, and the Army’s Scout helicopter.

Overall, the spike in Navy contracts for F-35s in 2013 obscured significant declines in numerous other Aircraft programs, and many of the other programs that saw significant increases in contract obligations in 2013 declined heavily in 2014. Data labeling is a significant issue in this category, as over $5.6 billion in contract obligations for rotary and fixed-wing in 2014 aircraft (representing one-seventh of total Aircraft contract obligations) are not properly classified under their parent programs.

**Electronics & Communications**

Contract obligations for Electronics & Communications (E&C) declined at the same rate (-37%) as overall DoD products between 2008 and 2014.

In 2013, E&C contract obligations (-15%) declined roughly in parallel with overall DoD products. The largest decline was in the unhelpfully vague category of “miscellaneous communications equipment,” which fell by $700 million. There were also declines of between $200 and $300 million in obligations in a variety of product types.

---

18 This category includes the only PSC for unmanned systems, called “Drones.” PSCs do not provide any further visibility into what type of unmanned system a particular contract is for; CSIS urges policy makers to break out the catch-all “Drones” code to provide more data granularity.
In 2014, E&C contract obligations fell by only 4%, less than a third of the rate of overall DoD products. No category within E&C saw a change of more than $250 million.

No particular product type (other than the catch-all “miscellaneous” category) saw a particularly noteworthy change in either year since 2012, though obligations for “electronic countermeasure & quick reaction equipment” have declined by nearly $700 million since 2012, and obligations for “night vision equipment” have declined by nearly $400 million in the same period. It is notable that Army contract obligations for E&C have fallen significantly more steeply (-28% in 2013, and -11% in 2014) than for the rest of the DoD. On the whole, however, E&C contracts have weathered sequestration and its aftermath relatively well.

**Engines & Power Plants**

Between 2008 and 2014, contract obligations for Engines & Power Plants (E&PP; -11%) declined at less than a third of the rate of overall DoD products.

In 2013, as overall DoD products contract obligations declined sharply, obligations for E&PP actually increased by 4%, making E&PP the only category to see an increase under sequestration. Though there was a $500 million decline in contract obligations for unspecified “gas turbines & jet engines – aircraft,” that was offset by a $1.1 billion increase in contract obligations for “nuclear reactors.”

Obligations for E&PP continued to be relatively preserved in 2014, declining by only 3%, less than a quarter of the rate of overall DoD products. There was a nearly $500 million increase in E&PP contract obligations related to the F-35 program, which is somewhat interesting because of how it lagged the spike in obligations for the actual planes in 2013. The study team believes that E&PP obligations related to the F-35 have only in 2014 started to be properly labeled as such, as F-35 obligations never exceeded $100 million prior to 2014.

Overall, E&PP contract obligations have been remarkably stable during sequestration and its aftermath, which is likely a function of how major E&PP contracts are tied to large, prominent platforms that previous CSIS analysis found were largely protected under sequestration (Berteau et al., 2014).

**Ground Vehicles**

Between 2008 and 2014, contract obligations for Ground Vehicles declined by a remarkable 87%. Even accounting for the fact that 2008 represented a nearly 50% spike in obligations (related to MRAP purchases) compared to 2007 and 2009, obligations have dropped by around 80% from 2007 and 2009 levels.

In 2013, contract obligations for ground vehicles declined by 41%, the largest fall for any product category under sequestration. The major drivers of this decline were a $1.8 billion decline in unspecified obligations for “combat assault & tactical vehicles” (primarily within the Marines), a $500 million decline in obligations for unspecified “combat assault & tactical vehicles – wheeled” (primarily within the Army), and a $400 million decline in obligations related to the Bradley Fighting Vehicle.

In 2014, by contrast, Ground Vehicles contract obligations (-16%) only declined slightly faster than did overall DoD products. The Army saw further reductions in obligations for unspecified “combat assault and tactical vehicles – wheeled” (-$400 million), while obligations for “trucks and truck tractors – wheeled” fell by $700 million. Obligations for unspecified “combat assault & tactical vehicles” increased by nearly $500 million, but whereas the large decline for this product type in 2013 was for the Marines, the increase was primarily in the Army.
DoD contract obligations for Ground Vehicles have cratered since their peak in 2008, and are lower in 2014 than they have been since 1999, right as the Army was beginning its ill-fated Future Combat Systems program that was intended to provide replacements for its aging fleet of ground vehicles. The inability of the Army since then to get a new ground vehicle program into full production, as well as the end of major combat operations in Iraq and Afghanistan, are the main drivers of this precipitous decline.

**Missiles & Space**

Contract obligations for Missiles & Space (M&S) declined by 18% between 2008 and 2014, less than half the rate of overall DoD products.

In both 2013 and 2014, contract obligations for M&S declined by 15%, roughly in parallel with the decline in overall DoD products. The decline in 2013 was almost entirely the result of a data coding change: $2.3 billion in obligations for space launches under the EELV program were reclassified as services rather than products. By contrast, 2014 saw what appear to be real, significant changes: a $2 billion decline in obligations for MDA Support for “guided missiles,” a $400 million decline related to the AMRAAM missile program, and notable increases in obligations related to the JASSM cruise missile program and the NAVSTAR GPS satellite program.

**Ships**

During the 2008–2014 budget downturn, as overall DoD products contract obligations declined by 37%, obligations for Ships actually increased by 37%, making Ships the only category to see an increase over this period.

In 2013, contract obligations for Ships declined by 10%, somewhat more slowly than for overall DoD products. As discussed in the Navy section earlier, there were notable increases in obligations related to the CVN-68 aircraft carrier, DDG-51 destroyer, and unspecified “combat ships and landing vessels,” along with significant declines in obligations related to the Ford-class aircraft carrier and the America-class amphibious assault ship.

As overall DoD products contract obligations declined by 14% in 2014, contract obligations for Ships actually increased by 7%; Ships was the only product category to see an increase in obligations in the year after sequestration. Obligations related to the Ford-class aircraft carrier and the SSN-74 Virginia-class submarine both nearly doubled in 2014.

The relative preservation of contract obligations for Ships in 2013, and the growth in 2014, is likely the result of a new policy development and a couple of existing factors. First, the “rebalance to the Asia-Pacific” has put a focus on the importance of sea platforms to future U.S. strategic interests and goals. Secondly, many of the major Ships programs are long-term production contracts, and a number are under multi-year procurement agreements; cutting or delaying funding for these programs would likely lead to greater costs over the long term.

**Final Thoughts**

This report presents only a fraction of the data compiled for this study; the final report coming out of this research effort will examine a broader set of contract characteristics, go into greater depth on some of the trends identified here, and focus more deeply on the non-MDAP-centric product categories. This report represents what the study team believes to be among the most notable and immediately relevant findings. The following are the key takeaways from this analysis:

- The overall rate of effective competition for products seems to have been largely unaffected both by the post–Cold War drawdown of the 1990s and the
current drawdown plus sequestration. There is a logic to the assumption that rates of competition would increase as the same pool of vendors fought for a declining pool of contract dollars, but that assumption has not been borne out in the data. By contrast, there have been notable declines in competition rates for MDAP-heavy product categories, such as Aircraft and Missiles & Space, particularly in the years after the Last Supper industry consolidation.

- The current downturn has seen the relative preservation of contract obligations going to the Big 6 defense vendors, despite the divestment of Northrop Grumman’s large shipbuilding unit into Huntington Ingalls Industries. This is likely a reflection of the preservation of the largest, most high-profile programs as budgets declined; these programs are disproportionately contracted to the Big 6.

- Sequestration has had an enormous impact on defense products contracting, even in the context of the overall decline since the peak in 2008. Overall defense products contract obligations have declined by 37% since 2008, but three-fifths of that decline occurred in just 2013 and 2014.

- Cuts in obligations were not evenly distributed among the major DoD components and product categories in 2013 and 2014. Many of the components and product categories that saw the most significant declines in 2013 were relatively preserved in 2014, and vice versa. This also applies to specific programs—many programs that saw major increases or cuts in obligations in 2013 saw significant reversals in 2014.

Data quality, both in current data and the pre-2000 data, remains a significant barrier to some areas of analysis. The major outstanding data quality issues include the following:

- The grouping of all unmanned systems contracts into a single PSC for “Drones” does not provide sufficient data granularity for an increasingly important segment of the DoD products portfolio.

- Data labeling issues within the Air Force, which show almost no obligations associated with the F-35 and the reuse of old codes for new projects, as shown by nearly $1.5 billion associated with a 1970s Army anti-tank missile program, are reason for concern and skepticism about the reliability of the valuable system equipment code field.

- The huge amount of contracts apparently either missing from or significantly undervalued in FPDS between 1990 and 1994 are a significant bar to any analysis trying to use FPDS data to examine the previous budget drawdown of the 1990s; this is a vexing issue that negatively impacts the researcher’s ability to distill out policy recommendations for policy makers.

Going forward, the study team will continue to work to find solutions to existing data quality issues, and to highlight those issues to policy makers where solutions are not possible on our end. CSIS will also continue to dig deeper into the wide range of data available on contract and vendor characteristics in the DoD products contracting market to identify key trends and possible lessons learned.

References

**Acknowledgments**

The study team would like to acknowledge the contributions of Jacob Bell, Gabriel Coll, Samantha Cohen, Ryan Crotty, Guy Nzeribe, Madison Riley, and John Vick to the quality and robustness of this research effort.
Contractor Past Performance Information: An Analysis of Assessment Narratives and Objective Ratings

Rene G. Rendon—is a nationally recognized authority in the areas of supply management, contract management, and project management. He is currently on the faculty of the U.S. Naval Postgraduate School (NPS), where he teaches in the MBA and Master of Science programs. Prior to his appointment at NPS, he served for more than 22 years as an acquisition and contracting officer in the U.S. Air Force, retiring at the rank of lieutenant colonel. His Air Force career included assignments as a warranted contracting officer for the Peacekeeper ICBM, Maverick Missile, C-20 (Gulfstream IV), and the F-22 Raptor. He was also a contracting squadron commander for an Air Force pilot training base and the director of contracting for the Air Force’s Space-based Infrared satellite system, and the Evolved Expendable Launch Vehicle rocket program.

Dr. Rendon has taught contract management courses for the UCLA Government Contracts program; he was also a senior faculty member for the Keller Graduate School of Management, where he taught MBA courses in project management and contract management. He is a graduate of the U.S. Air Force Squadron Officer School, Air Command and Staff College, Air War College, and the Department of Defense Systems Management College. Dr. Rendon is Level III certified in both program management and contracting under the Defense Acquisition Workforce Improvement Act (DAWIA) program. He is also a certified professional contracts manager (CPCM) with the National Contract Management Association (NCMA), a certified purchasing manager (C.P.M.) with the Institute for Supply Management (ISM), and a certified project management professional (PMP) with the Project Management Institute (PMI). He has received the prestigious Fellow Award from NCMA, and he was recognized with the U.S. Air Force Outstanding Officer in Contracting Award. He has also received the NCMA National Education Award and the NCMA Outstanding Fellow Award. Dr. Rendon is a member of the ISM Certification Committee as well as on the Editorial Review Board for the ISM Inside Supply Management magazine. He is a member of the NCMA Board of Advisors as well as associate editor for its Journal of Contract Management. Dr. Rendon’s publications include Government Contracting Basics (2007), U.S. Military Program Management: Lessons Learned & Best Practices (2007), and Contract Management Organizational Assessment Tools (2005). He has also published scholarly articles in Contract Management magazine, the Journal of Contract Management, Program Manager magazine, Project Management Journal, and PM Network magazine. He is a frequent speaker at universities and professional conferences and provides consulting to both government and industry. [rgrendon@nps.edu]

Uday Apte—is distinguished professor of Operations Management at the Graduate School of Business and Public Policy (GSBPP), Naval Postgraduate School (NPS), Monterey, CA. He also serves as the associate dean of Research and Development at GSBPP. Before joining NPS, Dr. Apte taught at the Wharton School, University of Pennsylvania, Philadelphia, and at the Cox School of Business, Southern Methodist University, Dallas. He is experienced in teaching a range of operations management and management science courses in the executive and full-time MBA programs. Prior to his career in academia, Dr. Apte worked for over 10 years managing operations and information systems in the financial services and utility industries. Since then, he has consulted with several major U.S. corporations and international organizations.

Areas of Dr. Apte’s research interests include managing service operations, supply chain management, technology management, and globalization of information-intensive services. He has completed over 10 sponsored research projects for the U.S. Department of Defense and has published over 60 articles, five of which have won awards from professional societies. His research articles have been published in prestigious journals including Management Science, Interfaces, Production and Operations Management (POM), Journal of Operations Management (JOM), Decision Sciences Journal (DSJ), IIE Transactions, and MIS Quarterly. He has co-authored two books, Manufacturing Automation and Managing in the Information Economy. Dr. Apte has served as a vice president of colleges at Production and Operations Management Society (POMS), as a founder and president of the POMS College of Service Operations, and as guest editor of POM journal. Currently he serves as the senior editor of POM and as associate editor of DSJ.
Dr. Apte holds a PhD in decision sciences from the Wharton School, University of Pennsylvania. His earlier academic background includes an MBA from the Asian Institute of Management, Manila, Philippines, and Bachelor of Technology (chemical engineering) from the Indian Institute of Technology, Bombay, India. [umapte@nps.edu]

Michael Dixon—is an assistant professor of Operations Management at the Graduate School of Business and Public Policy at the Naval Postgraduate School (NPS). He is interested in service operations management and specifically the effect of operational decisions and service design on customer experience. Dr. Dixon conducts management research from an analytical perspective and analyzes large archival data sets to model customer behavior, finding ways to represent the behavioral aspects of service design in quantitative and analytical ways. Specifically, he is interested in understanding the role of experiential sequence on participant perception. In past research, he analyzed season subscription ticket sales for a large performing arts venue and showed that season subscription re-purchases are influenced by event sequences from previous seasons leading to managerial direction for event planners to properly schedule performances across a season to maximize experiential effects. He teaches managerial statistics to MBA students at NPS and has been awarded the Liskin Teaching Award two times in the past two years for his teaching efforts.

Dr. Dixon is a co-editor of the Production and Operations Management Society (POMS) Chronicle newsletter and has been asked to serve as a session chair in Decision Sciences Conference in 2014. He has publications in the Journal of Operations Management and articles under review at the Production and Operations Management, Journal of Service Management, and Journal of Service Research. He has served as a reviewer for Management Science, Production and Operations Management, Journal of Operations Management, Journal of Service Research, Service Science, and the Decision Sciences Journal of Innovative Education. [mjdixon@nps.edu]

Abstract
The Department of Defense (DoD) acquires billions of dollars of supplies and services every year. In fiscal year (FY) 2013, the DoD obligated over $258 billion for military-unique weapon systems as well as commercial supplies and services. An integral part of the DoD’s contract management process is the source selection phase when offerors’ proposals are evaluated and the contract award decision is made. A critical aspect of the source selection phase is the evaluation of contractor past performance information as part of the overall proposal evaluation process. The DoD uses the Past Performance Information Retrieval System (PPIRS), which consists of contractor report cards extracted from the Contractor Performance Assessment Reporting System (CPARS). In this research, we examine the value of CPARS report card narratives for service contracts as they relate to their associated objective scores. Our primary focus in this research is to examine if the CPARS report card written narrative section provides value to the contractor performance evaluation process. Our data analysis includes sentiment and statistical analysis, as well as interviews with government agency contracting professionals. Using CPARS data, narrative analyses, and interviews, we answer the following research questions: (1) To what degree are government contracting professionals submitting to CPARS contractor performance narratives in accordance with the guidelines provided in the CPARS user’s manual? (2) What is the added value of the contractor performance narratives beyond the value of the objective scores for performance? (3) What is the statistical relationship between the sentiment contained in the narratives and the objective scores for contractor evaluations? The research revealed that there are a variety of opportunities to improve the contracting process specifically related to the narrative portion of past performance assessment reports.

Introduction
The Government Accountability Office (GAO) reported that the poor management of service contracts has undermined the government’s ability to obtain a good value for the money spent and has contributed to the GAO’s decision to designate management of services contracts as a high-risk area for the Department of Defense (DoD; GAO, 2013b). In
fact, as stressed in a recent memorandum for acquisition professionals by the under secretary of defense for acquisition, technology, and logistics (US[AT&L]), improving the efficiency of acquisition of products and services is of utmost importance to the DoD (USD [AT&L], 2010. More specifically, in a later memorandum, the USD (AT&L) focused on “improving tradecraft in services acquisition” (USD [AT&L], 2010, p. 5) by strengthening and improving the services contracting process. An important part of the services acquisition process is the evaluation of contractor past performance information using the Contractor Performance Assessment Reporting System (CPARS).

The CPARS report is initiated by DoD contracting officers during the contract closeout phase of the contract management process for documenting contractor performance information on the completed contract. It is also used by DoD contracting officers during the source selection phase as part of the evaluation of contractor proposals. The CPARS report contains contractor performance information using objective scores in five categories: Quality, Schedule, Cost Control, Business Relations, and Management of Key Personnel. In addition to these five objective categories, the CPARS reports also provide a subjective narrative section where the contracting officer provides a descriptive narrative of the contractor’s performance.

Although the use of contractor past performance information is an important aspect of the DoD contract management process, the GAO has identified many process deficiencies in the documentation and management of CPARS reports. GAO reports have shown that DoD agencies do not always complete the required contractor past performance reports (GAO, 2007, 2009b, 2013a, 2014). The 2012 National Defense Authorization Act (NDAA) required the DoD to “develop a strategy to ensure that evaluations in past performance databases used for making source selection decisions are complete, timely, and accurate” (GAO, 2014, p. 4). Additionally, the 2013 NDAA required a “government-wide strategy to ensure that timely, accurate, and complete information on contractor performance is included in past performance databases used by executive agencies for making source selection decisions” (GAO, 2014, p. 1).

Subsequently, the DoD increased focus on training and education for contracting officers, which resulted in an increase in contractor performance assessments being completed and submitted. In 2013, the GAO noted significant gains in CPARS completion rates: 56% of required reports were completed in 2011 while 74% were completed in 2013. However, according to the same GAO report, over half of the CPARS reports were submitted late. More importantly, many CPARS reports contain narratives that are either insufficiently detailed or are in conflict with their associated objective scores. Late reports lacking sufficient accurate information provide less-than-optimal information to the contracting professionals that rely on these report cards for source selection and contract administration purposes (GAO, 2013a).

The purpose of this research is to determine the value of the CPARS narratives in services acquisition by comparing the relationships between the subjective narratives and the associated objective scores. Our analysis allows us to suggest improvements to the CPARS management process, thus leading to greater and more effective utilization of the CPARS reports in services acquisition.

Research Methodology

This research examines the value of CPARS report card narratives for service contracts as they relate to their associated objective scores. Our primary focus in this research is to examine if the CPARS report card written narrative section provides value to the contractor performance evaluation process. Our data analysis includes sentiment and
statistical analysis, as well as interviews with government agency contracting professionals. Using CPARS data collected by graduate students, Wilhite, Stover, and Hart (2013), and narrative analyses and interviews conducted by graduate students Black, Henley, and Clute (2014), we answer the following research questions:

1. What is the added value of the contractor performance narratives beyond the value of the objective scores for performance?
2. What is the statistical relationship between the sentiment contained in the narratives and the objective scores for contractor evaluations?
3. To what degree do the interview findings contradict, support, or enhance the findings for our research questions?

Literature Review

Federal procurement policy requires that agencies collect information regarding a contractor’s performance under previously awarded contracts for all contracts over $100,000 and make that information available for use in future contract award decisions (Nash et al., 2007). The collection of contractor performance information occurs during the contract closeout phase using the DoD CPARS (Rendon & Snider, 2008).

The CPARS assessment data reflects the contractor’s performance in specific areas including quality, schedule, cost control, business relations, management of key personnel, and utilization of small business. The “Quality” rating assesses the contractor’s qualitative performance and compares it to the requirements stated in the contract. The “Schedule” rating assesses the contractor’s ability to meet schedules outlined in the contract such as milestones, task orders, delivery schedules, and administrative requirements. The “Cost Control” rating assesses the contractor’s ability to forecast, manage, and control the costs associated with performing contracted services. The “Business Relations” rating assesses the contractor’s ability to coordinate its business activities such as cooperate behavior, customer satisfaction, management, and attitude towards customers. The “Management of Key Personnel” rating assesses the contractor’s ability to maintain qualified individuals in key positions as outlined in the contract. The “Utilization of Small Business” rating assesses the contractor’s ability to integrate small businesses in the execution of the contract (Wilhite et al., 2013).

The CPARS assessment rates the contractor in these areas using the rating scales Exceptional, Very Good, Satisfactory, Marginal, and Unsatisfactory. It should be noted that the contractor is allowed to review the CPARS assessment and provide comments back to the government assessing official prior to the government finalizing the CPARS report.

During the source selection phase of government negotiated procurement, contractor performance information is used in evaluating offerors and in making a contract award decision (Rendon & Snider, 2008). In this phase, the government agency accesses the contractor performance information through the DoD Past Performance Information Retrieval System Report Cards (PPIRS-RC) database. During source selection in the evaluation of offeror’s proposals, the government agency uses the contractor past performance information to determine if the offeror meets the required standards of responsibility as stated in the federal procurement policy, and, depending on the basis of the award stipulated in the solicitation, uses the contractor’s past performance ratings to justify an award to a higher-priced offeror.

The contractor performance information reported in CPARS and accessible through PPIRS provides outcome-based data that can be used to identify successful contracts. The
successful contracts determined by using contractor performance information have been used in our previous research to identify the contract variables that lead to contract success.

In 2014, with the assistance of our graduate students (Wilhite et al., 2013), we accessed the past performance database to collect contractor performance ratings on 715 completed Army services contracts to determine if the contracts were successful or not successful. Using statistical analysis, we investigated whether certain contracting variables such as type of service, contract dollar value, level of competition, and contract type affected the success of the contract. The detailed results of our analysis are presented in Rendon, Apte, & Dixon (2014). Our research findings concluded that the S type services (Utilities and Housekeeping) had the highest failure rate of all the product service codes analyzed. We also found that contracts with a dollar value from $50 million to $1 billion had the highest failure rate of all the contract categories. We found that contracts competed competitively had the highest failure rate when compared to the other two forms of competition available. Furthermore, we found that contracts structured as a combination contract had the highest failure rate when compared to the other five types of available contracts. Finally, the results of our significance testing showed that Contractual Amounts and Contract Type were our only statistically significant variables (Wilhite et al., 2013).

Our past research using CPARS data identified some interesting areas worthy of further exploration. These areas include analyzing the narrative portion of the CPARS ratings to determine alignment with the objective ratings, as well as the value added, not only in the narrative portions, but also in the usefulness of the CPARS as a contractor assessment tool. This is the focus of our current research project.

**Research Design**

Our research examines the value of the CPARS report card narratives for service contracts as they relate to their associated objective scores. The primary focus in this research is examining if the CPARS report card written narrative section provides value to the contractor performance evaluation process. Our data analysis included a sentiment analysis and statistical analysis, as well as interviews with government agency contracting professionals.

With the assistance of our most recent MBA thesis students (Black, Henley, and Clute), we performed a sentiment analysis of the 715 Army service contract CPARS report card narratives accessed in our previous research (Rendon et al., 2014). Our students used the CPARS Quality Checklist as a basis for developing the criteria for the categories and values for the sentiment analysis (CPARS Best Practices, CPAR Quality Checklist, n.d.). In the sentiment analysis, the student researchers scored each narrative along the dimensions of quality, robustness, compliance with directions in the CPARS Quality Checklist, and its value and content compared to its related objective scores from the CPARS report cards. Independent researchers’ scores were compared across a small sample to ensure inter-rater reliability.

We conducted a statistical analysis of the relationship between the sentiment analysis scores and their associated objective rating scores. This analysis investigated correlating relationships between the sentiment scores and the objective rating scores for the same CPARS report. Our purpose was to explore the relationships between the sentiment scores and the objective rating scores to reveal the extent of the value of the narratives.

Our students (Black et al., 2014) also conducted interviews with contracting professionals from two DoD contracting agencies. These interviews focused on the
agencies’ use of CPARS and other sources of contractor past performance information and agencies’ value of the CPARS narratives compared to the objective rating scores.

Findings and Analysis

In this section, we present an analysis of our findings. The primary purpose of this research was to determine the value of the CPARS narratives in services acquisition by comparing the relationships between the subjective narratives and the objective scores. We first present the findings of the sentiment and statistical analysis by focusing on each of the criteria used in the analysis.

1. Do the narratives address all performance areas assessed? Overall, the narratives address all performance areas assessed ~82% of the time. This is less problematic with unsuccessful contracts at ~95% than with successful contracts at ~81%. The difference in the proportion of times that the narrative addresses all performance areas assessed in successful and unsuccessful contracts is statistically significant (p < .05; Black et al., 2014, p. 41).

2. Are narratives based on objective data? Overall, the narratives are based on objective data ~77% of the time. However, in unsuccessful contracts, the narratives are based on objective data 100% of the time. This is significantly different from the ~77% in successful contracts (p < .01; Black et al., 2014, p. 41).

3. Are narratives free of statements to avoid? Overall, the narratives are free of statements to avoid ~97% of the time. This is slightly more problematic with unsuccessful contracts at ~86% than with successful contracts at ~97% (p < .01; Black et al., 2014, p. 41).

4. Are narratives robust and comprehensive? Overall, the narratives are robust and comprehensive ~63% of the time. This is less problematic with unsuccessful contracts at ~91% than with successful contracts at ~62% (p < .01; Black et al., 2014, p. 41).

5. Could a layman understand the description of the work performed? Overall, the narratives are written so that a contracting layman should understand the work performed ~64% of the time. This is less problematic with unsuccessful contracts at ~82% than it is with successful contracts at ~64% (p < .05; Black et al., 2014, p. 41).

6. Is the narrative beneficial above and beyond objective scores? Using a Chi Square Test, we determined that there was a difference between successful and unsuccessful contracts in whether the narratives were beneficial above and beyond the objective scores. Unsuccessful contracts tended to have more beneficial CPARS report card narratives than successful contracts (Black et al., p. 42). Overall, the narrative provides an unsatisfactory amount of beneficial data to the user ~12% of the time. However, there were no unsuccessful contracts that provided an unsatisfactory amount of beneficial data. The narrative provides a marginal amount of beneficial data ~22% of the time. There were no unsuccessful contracts that provided a marginal amount of beneficial data. The narrative provides a satisfactory amount of beneficial data ~28% of the time. The narrative provides a very good amount of beneficial data ~21% of the time. The narrative provides an exceptional amount of beneficial data ~18% of the time. This is much more likely to occur with unsuccessful contracts than with successful contracts at ~17% (Black et al., 2014, p. 42).
7. Do the narratives correlate to the objective scores assigned? Using the Chi Square Test, we determined that there was not a difference between successful and unsuccessful contracts in whether the narrative correlates to the objective scores assigned. Overall, the narrative sentiment is contradictory to more than one of the objective scores assigned ~2% of the time. The narrative sentiment is contradictory to one of the objective scores assigned ~6% of the time. The narrative sentiment is satisfactory in describing accurately why the objective scores are assigned as they are ~28% of the time. The narrative sentiment is very successful in describing accurately why the objective scores are assigned as they are ~40% of the time. The narrative sentiment is exceptionally successful in describing accurately why the objective scores are assigned as they are ~24% of the time (Black et al., 2014, p. 42). Figure 1 summarizes the results of the statistical analysis.
Figure 1. Results of Statistical Database Analysis
(Black et al., 2014)

As previously discussed, our students also conducted interviews with contracting professionals from two DoD contracting agencies (Black et al., 2014). These interviews focused on the agencies' use of CPARS and other sources of contractor past performance information as well as these agencies' value of the CPARS narratives compared to the objective rating scores in the source selection process. The findings of these interviews are summarized as follows:

1. CPARS is still often not reliable, robust, or comprehensive enough. This results in source selection officials not placing a significant amount of weight on the past performance evaluation criteria (Black et al., 2014, p. 44).
2. Unsuccessful contracts tend to have more reliable, robust, and comprehensive past performance information available in their CPARS/PPIRS reports (Black et al., 2014, p. 45).

3. The appropriate amount of weight that should be assigned to the past performance evaluation criteria in making a source selection decision should be correlated to the source, availability, quality, and relevancy of the past performance information (Black et al., 2014, p. 45).

4. The information found in PPIRS sometimes contains information in the narrative that is either contradictory or does not quite match up with the objective scores. When the objective scores and narrative sentiment in PPIRS is mismatched, contracting professionals tend to give more weight to the narrative versus the objective scores (Black et al., 2014, p. 46).

5. Contracting professionals are not always applying due diligence in identifying the appropriate contractor entity (e.g., CAGE Code or DUNS number) in the CPARS reports. This results in contractor past performance information not being fully accessible in PPIRS (Black et al., 2014, p. 46).

6. There is a lack of reliable, robust, and comprehensive amount of past performance information available in PPIRS. This results in source selection officials soliciting contractors for references or asking contractors to fill out a past performance questionnaire (Black et al., 2014, p. 47).

7. The results of the interviews also identified recommendations for improving the quality of CPARS reports, incorporating data analytics tools into the PPIRS database, enhancing the monitoring of Contracting Officer Representative (COR) workload, improving acquisition workforce training on developing CPARS narratives, and improving the disclosure of CPARS program office audit results (Black et al., 2014, pp. 48–49).

Summary, Conclusions, and Recommendations

Summary

The DoD acquires billions of dollars of supplies and services every year. In FY 2013, the DoD obligated over $258 billion for military-unique weapon systems as well as commercial supplies and services (USA Spending, 2013). An integral part of the DoD’s contract management process is the source selection phase when offerors’ proposals are evaluated and the contract award decision is made. A critical aspect of the source selection phase is the evaluation of contractor past performance information as part of the overall proposal evaluation process. The DoD uses the Past Performance Information Retrieval Systems (PPIRS), which consists of contractor report cards extracted from the Contractor Performance Assessment Reporting System (CPARS). Although the use of contractor past performance information is an important aspect of the DoD contract management process, the GAO has identified many process deficiencies in the documentation and management of CPARS reports. GAO reports have shown that DoD agencies do not always complete the required contractor past performance reports (GAO, 2007, 2009b, 2013a, 2014). More importantly, many CPARS reports contain narratives that are either insufficiently detailed or conflict with their associated objective scores. Late reports lacking sufficient accurate information provide less-than-optimal information to the contracting professionals that rely on these report cards for source selection decisions.

The purpose of this research was to determine the value of the CPARS narratives in services acquisition by comparing the relationships between the subjective narratives and
the objective scores. Our primary focus in this research was examining if the CPARS report card written narrative section provides value to the contractor performance evaluation process. Our data analysis included sentiment and statistical analysis, as well as interviews with government agency contracting professionals.

Conclusions

Using CPARS data collected by graduate students Hart, Stover, and Wilhite, (2013), and narrative analyses and interviews conducted by graduate students Black, Henley, and Clute (2014), we answered the following research questions:

1. What is the added value of the contractor performance narratives beyond the value of the objective scores for performance? Contracting professionals are doing a better job at providing beneficial CPARS data in the narrative when the contract is unsuccessful versus when it is successful. Only 38.6% of the observed CPARS narratives, whether successful or unsuccessful, provided a very good or exceptional amount of beneficial data above and beyond what could be gleaned from looking over the objective scores assigned (Black et al., 2014, p. 51).

2. What is the statistical relationship between the sentiment contained in the narratives and the objective scores for contractor evaluations? Contracting professionals are developing CPARS narratives that contradict at least one of the objective scores assigned ~8.3% of the time. Contracting professionals were slightly better at matching the narrative sentiment to the objective scores in unsuccessful contracts (~81.8% scoring either very good or exceptional) than in successful contracts (~63.2% scoring either very good or exceptional; Black et al., 2014, p. 51).

3. To what degree do the interview findings contradict, support, or enhance the findings for our research questions? The results of the interviews found that the CPARS database is still often not reliable, robust, or comprehensive enough. We also found that unsuccessful contracts tend to have more reliable, robust, and comprehensive past performance information available in their CPARS/PPIRS reports. Additionally, the appropriate amount of weight that should be assigned to the past performance evaluation criteria in making a source selection decision should be correlated to the source, availability, quality, and relevancy of the past performance information. Our interviewees also stated that the information found in the PPIRS database sometimes contains information in the narrative that is either contradictory or does not quite match up with the objective scores. We also found that contracting professionals are not always applying due diligence in identifying the appropriate contractor entity in the CPARS reports, which is resulting in a lack of reliable, robust, and comprehensive amount of past performance information available in PPIRS. Finally, the interview results also identified recommendations for improving the quality of CPARS reports, incorporating data analytics tools into the PPIRS database, enhancing the monitoring of COR workload, improving acquisition workforce training on developing CPARS narratives, and improving the disclosure of CPARS program office audit results (Black et al., 2014, pp. 44–49).

Recommendations

Based on our conclusions, we identified the following five recommendations:
**Recommendation 1:** Training. Training should be implemented for all services acquisition members that interact with the CPARS and PPIRS databases. Training should focus on developing comprehensive narratives ensuring that acquisition team members can fully understand the work performed, address all performance areas assessed in their objective scores, and ensure the narratives are based on objective data (Black et al., 2014, pp. 54–55).

**Recommendation 2:** Process Improvement. The DoD needs to improve the quality of past performance report submissions in CPARS and PPIRS, improving the source, availability, quality, relevancy, and accuracy of the past performance information. This will allow acquisition teams to assign higher weights to past performance evaluation criteria in source selection decisions (Black et al., 2014, pp. 54–55).

**Recommendation 3:** Data Analytics. Additional data analysis tools should be incorporated into the CPARS and PPIRS database to better assist contracting professionals in identifying past performance trends for a particular contractor or specific type of service (Black et al., 2014, pp. 54–55).

**Recommendation 4:** Customer Feedback. The CPARS process should include customer feedback on contractor performance. Currently, only the acquisition team provides input to the CPARS report card. Customer input into CPARS will encourage the submission of more accurate and robust CPARS report cards (Black et al., 2014, pp. 54–55).

**Recommendation 5:** COR Manning Levels. Contracting Officer Representative (COR) manning levels should be reviewed throughout the DoD to ensure that organizations have sufficiently filled COR billets to manage the CPARS process (Black et al., 2014, pp. 54–55).

**References**


<table>
<thead>
<tr>
<th>Thursday, May 14, 2015</th>
<th>11:15 a.m. – 12:45 p.m.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chair: Gary Bliss</strong>, Director, Performance Assessments and Root Cause Analyses (PARCA), Office of the Assistant Secretary of Defense (Acquisition)</td>
<td></td>
</tr>
<tr>
<td><strong>Analytical Tools for Affordability Analysis</strong></td>
<td></td>
</tr>
<tr>
<td>David Tate, Institute for Defense Analyses</td>
<td></td>
</tr>
<tr>
<td><strong>A Dynamic Procurement Model for Assessing Competitive Options in the Context of Defense Acquisition</strong></td>
<td></td>
</tr>
<tr>
<td>Jennifer Lamping Lewis and Christopher Guo, RAND Corporation</td>
<td></td>
</tr>
<tr>
<td><strong>Government Contractors and Sticky SGA Costs</strong></td>
<td></td>
</tr>
<tr>
<td>Stephen Hansen, NPS</td>
<td></td>
</tr>
</tbody>
</table>

**Gary Bliss**—is the 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, 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 are 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, and 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 in mathematics and economics (highest honors in economics) from College of William and Mary.
Analytical Tools for Affordability Analysis

David Tate—joined the research staff of the Cost Analysis and Research Division at the Institute for Defense Analyses 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 the Johns Hopkins University, and MS and PhD degrees in operations research from Cornell University. [dtate@ida.org]

Abstract

Beginning with the Better Buying Power initiatives, the military components have been required to perform full life-cycle portfolio-level affordability analyses with respect to major defense acquisition programs at key milestones. Prior to this, there was no formal requirement to reconcile program acquisition baselines with resource forecasts beyond the five years of the Future Years Defense Program. This paper discusses the analytical challenges associated with Affordability Analysis and describes methods and tools that the Institute for Defense Analyses is developing to address these challenges.

Introduction

What Is Affordability?

Affordability means conducting a program at a cost constrained by the maximum resources the Department can allocate for that capability. Many of our programs flunk this basic test from their inception.

—Honorable Ashton B. Carter (2010), Under Secretary of Defense for Acquisition, Technology, and Logistics (USD[AT&L])

Affordability policy is about establishing the dollar amount the Component is willing to spend on the desired capability in the context of all other fiscal demands over the long term.

—Chad J. Ohlandt (2013)

Over the last two decades, the Department of Defense (DoD) has spent a historically unprecedented amount of money on acquisition programs that, in the end, did not deliver the warfighting capability they had been intended to provide. For example, the Army Acquisition Review commissioned by the Secretary of the Army and published in 2011 estimated that every year since 1996, the Army has spent more than $1B annually on programs that were ultimately cancelled. Since 2004, [35%–42%] per year of Army [Development, Testing, and Evaluation] funding has been lost to cancelled programs. (Decker & Wagner, 2011)

The Army is scarcely unique in this regard. All military components have seen significant program cancellations with little or nothing to show for the billions spent. In addition to cancellations, many other programs have been severely truncated or restructured, providing far less national security capability than had been envisioned.

While some programs have been canceled for technical reasons, or because the threat they were intended to counter went away, others have been canceled or curtailed...
simply because there was not enough money available to pay for everything. Affordability simply means having enough resources to be able to finish the programs you start.

What Is Affordability Analysis?

At the Milestone A Review [the DoD Component will] present an affordability analysis and proposed affordability goals based on the resources that are projected to be available to the DoD Component in the portfolio(s) or mission area(s) associated with the program under consideration. The analysis will be supported by a quantitative assessment of all of the programs in the prospective program’s portfolio or mission area that demonstrates the ability of the Component’s estimated budgets to fund the new program over its planned life cycle.

—DoD Instruction (DoDI) 5000.02, Operation of the Defense Acquisition System (USD[AT&L], 2015)

Affordability analysis is a DoD Component leadership responsibility that should involve the Component’s programming, resource planning, requirements, intelligence, and acquisition communities. The Department has a long history of starting programs that proved to be unaffordable. The result of this practice has been costly program cancelations and dramatic reductions in inventory objectives. Thus, the purpose of Affordability Analysis is to avoid starting or continuing programs that cannot be produced and supported within reasonable expectations for future budgets.

—DoDI 5000.02 (USD[AT&L], 2015, Enclosure 8)

Beginning with the Better Buying Power (BBP) memorandum and directive of 2010, and continuing through BBP 2.0, BBP 3.0, and the newly revised DoDI 5000.02, Operation of the Defense Acquisition System, the Office of the Secretary of Defense (OSD) has defined and implemented a series of new requirements for the military components to address affordability issues at the inception of new programs and at every subsequent milestone (DoD, n.d.)¹. The required Affordability Analysis should be performed at the portfolio level, over a planning horizon of decades. The guidance is explicit: the responsibility for Affordability Analysis lies with the component leadership, not with the program. Only the component leadership has the necessary understanding of component priorities, risk tolerance, and resource forecasts to support the required analysis, and only the component leadership has the authority to stretch, truncate, or cancel some programs in order to make room for others.

What Tools Do Affordability Analysts Need?

Predictive Costing—“What If?”

The first fundamental requirement for Affordability Analysis is the ability to consider hypothetical alternative futures in which the current plans cannot be executed. Current law requires that planned acquisition costs must fit within forecast budgets through the five-year

¹ Better Buying Power (BBP) 1.0 was introduced in 2010 as part of the DoD’s Efficiency Initiative. The stated objective was to deliver warfighting capabilities needed within the constraints of a declining defense budget by achieving better buying power for the warfighters and taxpayer.
Future Years Defense Program (FYDP), but not in the “outyears” after that. This leads to the phenomenon known as “the bow wave,” in which the total planned costs of all acquisition programs exceed reasonable budget projections in the years immediately beyond the FYDP. Each new President's Budget imposes fiscal reality on one additional year, pushing the excess off into the future. The new Affordability Analysis requirement extends the constraints of the FYDP out to a 30-year planning horizon, eliminating the bow wave. The top line budget assumptions for each component are provided by OSD.

To meet this requirement, component programmers need to be able to understand how annual program costs would change under hypothetical alternative schedules. As we shall see below, this is not an easy task.

**Consistency Checking**

Another important requirement for Affordability Analysis is the ability to assess the consistency between various affordability submissions. The new guidance requires Affordability Analyses to support milestone decisions for specific programs. These generally provide considerable detail about the program in question, some detail about other programs in the same acquisition portfolio, and minimal detail about other portfolios. Decision-makers need to be able to recognize whether new Affordability Analyses are consistent with past submissions by other programs and portfolios. They need to be able to distinguish updated information (e.g., new cost estimates or quantity requirements) from inconsistent assumptions. They also need to be able to understand the implications of reconciling inconsistent submissions using the most authoritative available data for all programs.

**Risk Modeling**

Because all Affordability Analyses are uncertain, it is essential for decision-makers (both within the components and in the oversight community) to understand the potential implications of that uncertainty. One important tool in this regard is sensitivity analysis—the ability to see the overall consequences of changing specific assumptions or forecasts. For affordability, it will be important to understand sensitivity to cost estimating assumptions, sensitivity to budget forecasts, and sensitivity to future needs for not-yet-defined systems.

More generally, “affordability” is not a yes/no question. It makes much more sense to think of the probability that a given portfolio of programs can be acquired within a given budget, even though both program cost estimates and budget forecasts are subject to unknown errors. An ideal tool for Affordability Analysis will support probabilistic modeling of individual program outcomes, portfolio contents, and budget forecasts.

**Visualization**

Finally, it is important for any Affordability Analysis support tool to provide outputs that are informative and understandable. Numerical reporting is important, but data visualization can greatly enhance the utility of any analytical tool. In particular, dynamic visualizations that highlight the patterns of change that occur as budgets grow tighter (or portfolio contents grow) would be very useful. To support this requirement, it will be necessary for the tool to be able to perform its “What if?” computations in near-real time.

**What Are the Analytical Challenges of Affordability Analysis?**

Why is Affordability Analysis hard? Setting aside the political challenges of competing interests within a given military component, it might not be obvious that there are also substantive technical challenges in figuring out whether or not a set of proposed programs is affordable. In this section, we discuss three necessary enablers for effective
Affordability Analysis and the technical challenges associated with each. All three of these challenges arise directly from the nature of program cost estimates.

Any Affordability Analysis must begin with cost estimates for the various programs in the relevant portfolio. These estimates generally take the form of annual quantity and cost forecasts that reflect the program offices’ planned (or proposed) development and fielding schedules. If the sum of these forecasts fits within the predicted available budget for the portfolio, then the portfolio would seem to be affordable. But what if it does not fit? And what about the possibility that the budget forecast or the cost estimate (or both) might turn out to have been overly optimistic? To address those concerns, we need to be able to predict the annual costs for production schedules other than the planned schedule—and that turns out to be rather tricky.

**Estimating the Annual Costs of Hypothetical Schedules**

The problem is easy to state: “The program says that the planned schedule of lot quantities would result in these estimates of annual lot costs. How much would each lot cost if we bought them according to this alternative schedule of procurement quantities instead?”

The DoD has detailed records of cost and schedule for hundreds of acquisition programs, going back decades. Given that wealth of data, why is it hard to figure out how schedule affects cost? There are several important analytical obstacles.

**Obstacle #1: Cause Versus Effect**

Consider three procurement programs: A, B, and C. Program A is doing fine, but due to overall budget reductions, its production schedule is going to be stretched, which will increase unit costs. Program B has just announced significant cost growth—not caused by a schedule change—that has made its planned production schedule too expensive, given other priorities in the portfolio. Program C has been experiencing integration issues—its electronics are going to require a new design that uses a more expensive subcomponent, which will have to be retrofitted into existing units. This means both a cost increase (due to the new component) and a schedule slip (to accommodate the new design and the rework).

For all of these programs, unit cost went up and average production rate went down. Causally, though, we have three distinct cases:

- For Program A, schedule stretch caused cost growth.
- For Program B, cost growth caused schedule stretch.
- For Program C, technical issues caused both cost and schedule growth.

Although we are trying to understand only the first of those mechanisms, we cannot tell just by looking at historical numbers which case was in effect—or whether it was some mix of all of them—for a given program. We need a way to isolate the Program A effect from the others.

**Obstacle #2: Not All Costs React the Same Way to Schedule Changes**

Since 2006, Selected Acquisition Reports have broken out cost forecasts into subcategories: end-item recurring flyaway costs, non–end-item recurring flyaway costs, nonrecurring costs, and two categories of support costs. This is very helpful, because we do not expect all of those costs to react identically to a change in production schedule. End-item recurring costs should be most directly affected, while nonrecurring costs and non-spare support might not be affected at all. Any econometric model of how schedule affects cost should take advantage of these different cost categories and treat them separately when they are known.
Obstacle #3: Limited Relevant Data

Even if we have multiple historical cost estimates for Program A, those past estimates might not tell us anything about the relationship today between Program A’s production schedule and lot costs. Since those past estimates were developed, Program A may have changed in any number of ways—new designs, revised cost estimates, requirements changes, technology insertions, planned product improvements, new contracts, new demands from the field, and so forth.

Unless we could somehow correct for all of the program changes other than schedule, those past forecasts do not tell us what the estimated cost would be today for that previous planned schedule. More generally, we seldom get to see multiple proposed schedules (and their associated costs) for the same exact program. Since we are trying to figure out how cost varies as schedule varies, this is a major limitation.

Obstacle #4: Rate Effects and Learning Curves

The cost of producing one unit of a product depends on the production rate. For one thing, at lower production rates, the indirect costs (overhead) of the producer and the fixed costs of production get amortized over fewer units, so that each unit bears a higher proportion of those costs. In addition, there can be logistical issues related to supply chains and staffing that make it inefficient to produce at low rates. We will discuss the mechanisms of rate effects below; for the moment it suffices to note that they exist. As a result, the cost of a unit depends on how many other units you make in the same year.

In addition, even at constant production rates, unit costs are generally not constant over the production life of a program. In particular, most procurement programs exhibit learning curves, in which the marginal cost of successive units decreases as a function of cumulative production. As a result, the cost of a given unit depends not only on how many other units you make in that year, but which units (cumulative) they are. In the standard model, the cost of the nth unit produced is predicted to be $T_1 n^\beta$, where $T_1$ is a notional first unit cost and $\beta$ is a parameter that describes how quickly unit costs decline with cumulative production (Lee, 1997). This formulation is often expressed in terms of the learning “slope,” $S = 2^\beta$, where unit cost decreases by a factor of $S$ every time cumulative production doubles.

Shifting units of planned production from one year to another thus changes both the production rates in those years and the portion of the learning curve that falls in those years, resulting in nonlinear changes in annual costs.

Mechanisms and Models for Schedule Effects

There are several competing theories about how and why unit costs change when schedules change. These theories are not necessarily mutually exclusive, which makes it even trickier to figure out how to combine them into a coherent model.

Fixed Costs and Sticky Costs

As noted above, most of the indirect costs and some of the direct costs of producing a weapon system are incurred per unit time, rather than per unit produced. For example, the costs associated with running the program office do not depend much on the current production rate, or on how many units have been produced so far. Similarly, the indirect costs associated with contractor overhead are only a little bit sensitive to production rates. What’s more, overhead rates tend to be “sticky”—they don’t generally adjust instantaneously to changes in work level. A useful model of how cost depends on schedule should be able to
distinguish rate-insensitive costs from rate-sensitive costs, and estimate how sticky the fixed costs are.

**Learning and Forgetting**

It is not uncommon for unit production costs to follow a standard learning curve for most of the life of a program, but then start to climb upward again toward the end of the program. To account for this, C. Lanier Benkard (2000) suggested that producers improve in efficiency by gaining “experience” making units (learning), but that this experience depreciates at a constant rate (forgetting). Thus, early in production (when cumulative quantity is doubling frequently), or at high production rates (when more experience is being gained per unit time), learning behavior dominates. Late in the production run, or at low production rates, the gains from learning are visibly offset by forgetting.

We investigated this model, and found that it fits many historical programs quite well. It can also be improved by combining it with a fixed cost model, so that indirect and rate-insensitive costs are modeled separately, while rate-sensitive costs are modeled by a combination of learning and forgetting.

**Regulatory Lag**

Finally, William Rogerson (1994) has proposed that the interaction between unit cost reduction and production rate can be understood by looking at the incentives inherent in how procurement contracts are awarded. In general, a new fixed-price procurement contract is awarded for each annual lot, with a negotiated unit price based on the contractor’s demonstrated historical costs. A contractor who invests in management or tooling changes that reduce production costs will only realize extra profits from that investment until a new price is negotiated—typically two or three years later.

At high production rates, contractors have more incentive to invest in reducing production costs, because they will realize extra profits on many units during the two- to three-year “regulatory lag” period before the price is renegotiated downward to reflect the lower production costs. As a result, more potential cost-reducing investments will be cost-effective for them, given the need to make back the initial investment costs through higher profits.

Conversely, at lower production rates, contractors have less incentive to reduce costs, as well as fewer available cost-reduction alternatives that would provide the necessary return on investment. If this theory is correct, we should expect to see less learning at low production rates, and more learning at higher production rates. This is very different from the standard learning model, which assumes that the learning curve slope is an intrinsic characteristic of the system being produced and does not vary with rate.

It is possible to formulate an econometric model based on Rogerson’s (1994) framework, assuming a profit-maximizing contractor with a notional pool of cost-reducing investment options. We can also combine this model with a fixed-cost model, as we did with the learning-and-forgetting model. In theory, we could add forgetting to this model as well, but there are serious limits to how many model parameters can be estimated from the available data.

**Estimating the Impact on Schedules of Hypothetical Budgets**

In order to understand the impact of portfolio budgets on cost and schedule, it is not enough to be able to predict the annual costs associated with a given sequence of lot quantities. We also need to be able to predict how each program’s production schedule might adjust to accommodate a budget that is too tight for each program to execute its current individual plan.
In a prior research effort (Weber et al., 2003), we used optimization to find minimum-cost feasible schedules within a given budget, subject to constraints on minimum and maximum production rates and latest permitted delivery date. This approach is unsatisfactory for several reasons. First, the nonlinear mixed-integer optimization problem is quite difficult, requiring industrial-grade optimization tools and considerable time. Second, the optimized schedules are generally implausible—they rely on the portfolio manager’s ability to make trades among future years in a manner that is often politically impossible. They also tend to oscillate between minimum and maximum rates for a given program unless additional constraints or penalties are imposed. Finally, they give a very misleading view of the consequences of a change in forecast budget, since the bulk of the cost difference between the current planned schedule and the revised schedule is due to optimization of the current plan, not due to the change in budget.

For the current effort, we have chosen to use a greedy constructive heuristic to model the potential impact of a budget change on production plans within a portfolio. The heuristic attempts to maintain the same proportionate cumulative funding by program as in the plan, but is constrained by both the budget and the minimum (and maximum) feasible production rates of the various programs. The combination of fixed costs and minimum sustaining production rates implies that it may not be possible to match the planned relative funding levels very closely. This is especially true if some programs are near their minimum sustaining rates in the original plan. Those programs cannot be stretched further, so that the cost burden of stretching to fit under the budget will be borne disproportionately by the other programs. Optionally, the heuristic can also incorporate user-specified program priorities.

At present, the heuristic simply reports infeasibility if there is insufficient budget to produce each program at its minimum sustaining rate. Future versions may suggest quantity reductions where needed, based on user-provided priorities among programs and minimum useful quantities to field.

**Affordability Risk Analysis**

All of the analytical challenges discussed thus far make the implicit assumption that the estimated costs for each program in the portfolio are accurate. History teaches that this is not generally true; weapon system acquisition is notorious for cost growth. Without getting into the thorny question of what causes cost growth, we can nevertheless ask what implications cost growth has for Affordability Analysis.

If we knew both program costs and future budgets with certainty, affordability would be a yes/no question. Budget uncertainty can be handled to some extent through sensitivity analysis—how much do costs change as a function of future budget levels, and at what point does the portfolio become infeasible unless we reduce quantities or cancel programs? There is also a time dependency—one bottleneck year of low budget can render a set of programs infeasible, even if the overall funding level over the planning horizon is generally high.

Cost uncertainty is even trickier, because we need to know not only how the costs might change, but why they might change. Is the problem higher labor rates than expected? Higher fixed costs? Less learning? Faster forgetting? Some combination of all of those? The answers to those questions will affect how much further cost growth will be incurred when we realign production schedules to fit under the budget.

Because of the complexity of how cost depends on various econometric parameters and on schedule, there is little prospect of an analytical model of affordability risk. Even if we were confident that we knew the probability distributions for future budgets and for procurement cost growth in each program, that is not enough information to allow us to
derive the conditional distribution of resulting costs when all programs are modified to fit within a given budget. If we also want to be able to model uncertainty due to new programs arriving into the portfolio, the problem is even worse.

Monte Carlo methods seem like the most promising approach here. If we can characterize the cost risk in terms of probability distributions on the parameters of our econometric model, we can then simulate the future repeatedly to estimate the probability of affordability, the expected unit cost, the probability of a Nunn-McCurdy breach, the expected time to deliver the Nth unit, and other measures of interest. We can do this for a fixed budget, or we can allow the budget to vary randomly at the same time, including making room for future programs. The trick, of course, is deriving plausible probability models for the econometric parameters and budgets, based on historical data.

**APASS—The Acquisition Portfolio Affordability Support System**

We are currently developing a software environment for Affordability Analysis for our sponsors in the Office of the Secretary of Defense for Acquisition, Technology and Logistics (OSD[AT&L]) Acquisition Resources & Analysis, which we call the Acquisition Portfolio Affordability Support System, or APASS. The purpose of APASS is to provide an environment for evaluating and tracking the affordability of portfolios over time. APASS features are focused on five specific areas of support: tracking, forecasting, risk analysis, reporting, and visualization.

**Tracking**

APASS will provide both a “living” operational view and a historical record of portfolio affordability and cost and schedule data. This combination will support both current decision-making and assessment of historical trends and changes.

Affordability analyses are generally performed on behalf of individual programs for use at milestone reviews. These analyses generally feature detailed cost and schedule information for the program under review, aggregated cost and schedule information for other programs in that portfolio, and at best top-line summary data for other portfolios. APASS will allow managers to remember the information submitted at prior reviews, compare that information against the new information, and identify any inconsistencies or changes. In particular, APASS will support comparison of data sources with slightly different portfolio definitions. APASS will also be able to compare new submissions against routine DoD-wide data submissions, such as the annual Selected Acquisition Reports (SARs) or President’s Budget justification exhibits.

APASS will also allow analysts to merge multiple submissions into user-defined data sets. Authoritative data from multiple sources could thus be combined to create a “single integrated affordability picture” across all programs and portfolios. This would let managers maintain a living representation of the current best forecast of the acquisition future, with which they could explore the implications on that future of alternative budgets, new programs, or cost growth.

**Forecasting**

The primary analytical function of APASS is to predict the annual quantities and costs for all programs in a portfolio, conditioned on a portfolio budget other than the current planned budget. This function relies heavily on both the econometric modeling and the schedule adjustment heuristic described above.

Figure 1 shows an APASS screenshot of a notional current production plan for a set of portfolios. FYDP years are shown in darker hues; outyears are lighter shades of the same
color. If we drill down into a single portfolio, we can see the individual program plans and the projected portfolio budget (Figure 2). The combined plans exceed the projected budget, so we invoke the “fit to budget” heuristic to extrapolate how the portfolio might adapt to the lower budget. The result is shown in Figure 3. In this example, the total procurement cost for the portfolio increases significantly. The percent increase in unit cost is not shared evenly across programs; those programs that finish quickly or that are already near their minimum production rates will not be as strongly affected as those that have more flexibility to stretch.
This is the fundamental analytical tool in APASS; all other analyses rely on this ability to extrapolate costs and schedules in response to different levels of available funding. The cost forecasts produced by the econometric models and extrapolation heuristic are not “budget quality” cost estimates, but they provide credible extrapolations of the type and magnitude of cost and schedule impact that would be implied by alternative scenarios.

Risk Analysis

The initial implementation of APASS will focus on sensitivity analysis, particularly with respect to changes in budget or portfolio contents. It is particularly important for decision-makers to understand how program costs and delivery schedules might change under various budget scenarios, or in the event that a new high-priority program is added to a portfolio. It will also be possible to use APASS to assess the sensitivity of all programs to cost growth in a particular program.

Future versions of APASS will include Monte Carlo modeling capabilities, to simulate the potential consequences of multiple simultaneous sources of uncertainty. The output from these simulations would include probability distributions on unit cost and delivery timelines for individual programs, as well as correlations among programs within a portfolio.

Reporting

The analytical capabilities of APASS will only be useful if they are able to provide users with the information they need to support acquisition decisions and oversight. To that end, APASS will include both standard and customizable reporting of both cost and schedule baselines and derived affordability data.

Because APASS incorporates data from multiple sources, including speculative analyses, its reporting requirements will be somewhat more varied than those of typical acquisition data repositories. In particular, APASS will need to be able to merge conflicting data sources to produce “best guess” forecasts that draw individual program forecasts (and their corresponding econometric parametrizations) from separate sources, using the current most authoritative source for each. This merged forecast will typically differ from each of its source submissions in some ways. It will therefore be particularly important for APASS to
produce reports that highlight the differences between forecasts of the same program(s) from different sources. These comparison reports will also be useful for describing the unit cost and schedule impacts of various possible future portfolio-level budgets, as extrapolated by the schedule adjustment heuristic.

**Visualization**

As a complement to numerical reporting, APASS will also provide graphical visualizations of various baseline data and comparisons. To begin with, APASS will be able to produce the various graphical formats for presenting Affordability Analysis information that are recommended in the *Defense Acquisition Guidebook*, as revised in June 2013. APASS will also provide the ability to view animations of the sensitivity of a given portfolio to program cost growth, changes in budget level, or introduction of additional programs. Coupled animations, showing both the overall “sand chart” of program spending over time and an aligned display of the absolute or relative change from the baseline by program, will allow analysts to grasp the implications of alternative scenarios much more quickly than when using individual graphics or numerical displays.

**Summary and Conclusions**

Program life-cycle affordability is a cornerstone of DoD acquisition planning. Ultimately, the goal of Affordability Analysis is to supply decision-makers with the best available information and analysis about defense acquisition programs so that they can

- allocate defense resources efficiently and effectively, both within and across programs;
- consider the full range of cost and schedule alternatives;
- understand and manage acquisition risk at the program, portfolio, and Service level; and
- avoid the time and money wasted by starting programs that cannot be completed.

Affordability is the measure that allows decision-makers to allocate appropriate resources to their future operational requirements. Affordability is not a “yes or no” attribute; it is the degree to which uncertain future resources can be expected to permit execution of current and future programs. This degree is best characterized in terms of probabilities and sensitivity to deviations from current plans.

The Better Buying Power (BBP) memoranda and the newly-revised DoDI 5000.02, *Operation of the Defense Acquisition System*, from the OSD have defined and implemented new requirements for the military components to address affordability issues at the inception of new programs and at every subsequent milestone. This Affordability Analysis is to be performed at the portfolio level, over a planning horizon of decades. The responsibility is given to component leadership.

In order to plan their acquisition strategies, the components need to be able to assess affordability over a wide range of future scenarios without needing to return to the program offices or the component cost estimators for new cost estimates every time a different portfolio plan is considered. At the same time, OSD oversight organizations need to be able to reconcile current affordability analyses with past analyses, assess the sensitivity of current forecasts to potential disruptions, and characterize the level of risk inherent in current plans, in order to fulfill their oversight responsibilities with respect to acquisition.

APASS is being implemented as a tool to support both Affordability Analysis as required by current DoD regulations and OSD oversight of affordability planning across the
military components. To these ends, APASS will combine econometric modeling of procurement costs as a function of schedule, extrapolation of the effects of alternative budgets on portfolios of programs, and fusion of data from multiple divergent sources. APASS will also provide the reporting and visualizations services necessary to make these analytical capabilities practically useful on a day-to-day basis.

References
Abstract

Federal contractors must deal with an exceptional amount of paperwork and bureaucracy relative to firms that deal only with the private sector. I investigate whether federal contractors’ costs have different responses to revenue increases and decreases. I start by generating a set of federal focus firms that have a business unit name that incorporates the words federal, military, and defense. These firms have built their organizational structure around federal contracting. Because extra paperwork costs are likely to be part of the Selling, General, and Administrative (SGA) costs, I estimate a model of SGA sticky costs. I find that when revenues increase, federal focus firms have greater increases in SGA costs compared to controls. This increase is consistent with higher fulfillment costs for federal contracts. When revenues decrease, federal focus firms have a much lower decrease in SGA costs compared to controls. Federal focus firms have extremely sticky SGA costs. This stickiness is consistent with federal focus firms having higher fixed costs in their procurement systems.

Introduction

The Department of Defense (DoD) is currently implementing the Better Buying Power initiative (DoD, n.d.), which focuses on “the implementation of best practices to strengthen the Defense Department’s buying power, improve industry productivity, and provide an affordable, value-added military capability to the Warfighter.” An important plank of this approach is “Eliminate Unproductive Processes and Bureaucracy” (DoD, n.d.).

Government contractors have the normal concerns about costs of production/profit margin. However, they also must deal with an exceptional amount of paperwork relative to non-governmental contracts, and usually must hire specialized staff to generate and maintain the required information (Kovacic, 1992). The additional paperwork and extra staff clearly qualify as potentially unproductive processes and as bureaucracy.

My research examines whether these additional costs are sufficiently large to affect the firm’s financial statements. If these government-specific costs are real and substantial, then they should skew the cost behavior of government contractors relative to private sector firms doing comparable work.

I investigate the effect of these costs on the Income Statement. Income Statements have two major cost categories. The Cost of Goods Sold captures the product cost for units sold and is obtained from matching the cost of products to the units sold. These costs are fairly direct and have comparatively little wiggle room for adjustments. The Selling, General, and Administrative (SGA) costs reflect the marketing, administrative, and general overhead.
costs of the organization. These costs contain many allocations and are the likely place where any additional government contracting costs are going to show up.\(^3\)

My research builds upon prior work that investigates the behavior of SGA costs. A classic paper, Andersen, Banker, & Janakiraman (2003), investigated the behavior of SGA costs when revenue increased versus revenue decreased. They found that SGA costs are sticky in that SGA costs increase more when revenues rise than SGA costs decrease when revenues fall. Potential explanations for the stickiness include the existence of SGA fixed costs (Balakrishnan, Labro, & Soderstrom, 2010), or that managers are reluctant to reduce SGA capacity when they believe that a short run cut in revenues is likely to be reversed in the near future (Andersen et al., 2007; Balakrishnan & Gruca, 2008).

I extend the SGA sticky cost framework to investigate whether firms with significant federal contracting have different SGA responses to revenue increases and/or decreases. If a firm believes that the expertise in federal procurement practice is a core competency, then it will be reluctant to reduce its procurement staff in response to lower revenues. I conjecture that federal focus firms will have much stickier costs than will firms in the private sector.

An important innovation in my paper is the creation of a federal focus sample. Prior work has investigated the behavior of federal contractors by assuming that the firms with the largest dollar sales to the federal government are most affected by federal contracts (Wang & San Miguel, 2012). However, an important problem with this approach is that a very large firm like Proctor and Gamble may be one of the most important sellers to the federal government, but the federal sales may be a small percentage of Proctor and Gamble’s total sales. In addition, Proctor and Gamble may supply generic products such as toothpaste which do not require satisfying unique federal requirements (and related paperwork). For these reasons, it may not be correct to say that Proctor and Gamble is a federal focus firm.

My federal focus sample is created in a different fashion. I use Compustat’s Segment database to identify firms that have a business segment which contain the words federal, government, or military. These federal focus companies believe that federal business is so important that they have built their organization structure on this business. Their organization structure signals that they have a federal focus. While there is some overlap between my list of federal focus companies and the list of firms with the greatest sales to the federal government, roughly two thirds of my federal focus firms are not on the list of the largest 500 federal dollar contractors.

Once I generate the federal focus sample, I estimate an SGA sticky cost regression where I include interaction terms for federal focus firms. The results show that the control firms have mild SGA cost stickiness. Raising revenue by 1% leads to a 0.69% rise in SGA costs, while lowering revenues by 1% leads to a 0.63% decline in SGA costs. The main analysis compares the federal focus firms to the controls. There are significant differences for federal focus firms. Raising revenues by 1% leads to a 0.83% rise in SGA costs, a 20% increase over the control firms. Federal focus firms have higher ramp up in SGA costs, possibly due to increased paperwork for fulfilling government orders. In the same vein,

---

\(^3\) While the government has explicit rules about how these costs are presented in their forms, the published financial statements follow financial accounting rules known as the Generally Accepted Accounting Principles.
lowering revenues by 1% leads to a 0.45% drop in SGA costs, a 29% smaller drop than the
controls. SGA costs for federal focus companies are much stickier than for the control firms.
This increased stickiness could be due to a reluctance to fire highly trained procurement
staff or, alternatively, much higher fixed SGA costs.

Robustness tests suggest that my results are knife edge. Changing the time period
or the control group leads to insignificant results.

The remainder of the paper is organized as follows: The second section discusses
prior research, followed by a presentation of the sticky cost model. The sample is then
created and descriptive statistics are calculated. Estimation results follow, and the final
section contains the conclusion.

**Literature Review**

There has been a fair amount of prior work that has used published financial
statements to examine various features of acquisition and contracting. Arnold, McNicol, and
Fasana (2009) investigated the impact of various contract forms on contract performance.
Berteau, Levy, Ben-Ari, and Moore (2011) used financial statements to analyze the ability of
government contractors to obtain capital throughout defense booms and busts. Wang and
San Miguel (2012) investigated whether government contractors are obtaining excessive
profits. My work complements those prior works by providing an analysis of costs, an
important component in determining both performance and profits.

My work investigates the total cost numbers for the organization. The reason is
straightforward. Prior work has shown that firms have the incentives and capability to shift
costs from the private sector to the government sector (McGowan & Vendrzyk, 2002;
Rogerson, 1992). Cost shifting between the firm’s private and public units generates
canceling positive and negative entries when the total firm costs are calculated.

Balakrishnan and Gruca (2008) examined sticky costs at the department level in
hospitals. They found that in downturns, hospital administrators are reluctant to trim costs
and capacity in core activities directly related to patient care. Administrators first adjust costs
and capacity in peripheral areas. This research suggests that in a downturn, firms with a
federal focus may wish to maintain critical government contract-related processes.

**The Sticky Cost Model**

Andersen, Banker, and Janakiraman (ABJ; 2003) performed the seminal analysis of
SGA cost stickiness. Their model discriminated between periods when revenues increase
and those when revenues decrease. Costs are sticky when the costs have a greater rise
when revenues increase than costs fall when revenues decrease.

Equation 1 provides their basic sticky cost model specification,\(^4\)

\[
\log\left(\frac{SGA_{it}}{SGA_{i,t-1}}\right) = \alpha_0 + \alpha_1 \log\left(\frac{Revenue_{it}}{Revenue_{i,t-1}}\right) + \alpha_2 \text{DecrDum}_{i,t} \times \log\left(\frac{Revenue_{it}}{Revenue_{i,t-1}}\right) + \varepsilon_{i,t},
\]

\(^4\) The ABJ model allows for a cross-section analysis across a wide range of industries, with large
differences in the size of firm. Prior work (Davidson & MacKinnon, 1981) rejected a linear form in
favor of the log-log specification.
where \( DeccDum \) is one for firm \( i \) when sales revenues fall from period \( t-1 \) to \( t \) and is zero otherwise. If SGA costs are sticky, then the coefficient \( a2 \) should be negative and significant.

I extend Equation 1 to investigate whether federal focus firms have different SGA stickiness than control firms. I create a set of interaction variables which separate out the incremental effect of federal focus firms. Specifically, I use a dummy variable \( FSeg \), which is one if the firm has a separate government/federal/military segment and is zero otherwise. My SGA Sticky Cost model is presented in Equation 2.5

\[
\log \left( \frac{SGA_{it}}{SGA_{it-1}} \right) = \alpha_0 + \alpha_1 \log \left( \frac{Revenue_{it}}{Revenue_{it-1}} \right) + \alpha_2 DeccDum_{it} \log \left( \frac{Revenue_{it}}{Revenue_{it-1}} \right) \\
+ \alpha_3 FSeg + \alpha_4 FSeg * \log \left( \frac{Revenue_{it}}{Revenue_{it-1}} \right) \\
+ \alpha_5 FSeg * DeccDum_{it} * \log \left( \frac{Revenue_{it}}{Revenue_{it-1}} \right) + \delta_{it}.
\]  

(2)

While I believe that federal focus firms will have different SGA responses than the control firms, the direction of these changes is not obvious. If federal focus firms respond differently to revenue increases, then coefficient \( a_4 \) should be significant. If it is significantly positive (negative), then SGA costs have greater increases (decreases) than the controls. If federal focus firms respond differently to revenue decreases, then the expression \((a_4 + a_5)\) should be significant. If the sum is significantly positive (negative), then federal focus firms have less (more) sticky SGA costs than the controls.

Now that I have presented the model, I turn to the data.

Sample Creation and Descriptive Statistics

My sample was created in two steps. I first obtained a set of federal focus firms, then generated control firms for these companies.

Federal Focus Firms

The critical element in my research is identifying companies with a federal focus. Prior research (Wang & San Miguel, 2012) has examined the behavior of companies with the greatest dollar value contracts with the federal government. One problem with this approach is that a large dollar value may not reflect a federal focus. For instance, Proctor and Gamble has significant sales to the government, but its sales tend to be for off-the-shelf items such as toothpaste. Proctor and Gamble may not need to deal with issues involving government-specific specifications and may have minimal incremental paperwork requirements. Because of these issues, I use a different approach to identify federal focus firms.

Financial accounting standards require publicly traded firms to separately report information about major business segments. Segment reporting is intended to give

---

5 The variable \( DeccDum \) is defined in the text. The data for the other variables is drawn from Compustat, Fundamental Annual. Revenue is the Compustat variable sale, while SGA is the Compustat variable xsga.
information to investors and creditors regarding the financial results and position of the most important operating units of a company. Firms that report a federal segment have identified themselves as having a significant line of business related to the federal government.

Table 1, Panel A describes how I generated my list of federal focus firms. I began with all observations on the Compustat Segment database which lists all reported segments for all publicly traded companies in the United States. The Compustat Annual Updates—Segment database is comparatively new and only has data for the last four years, 2010–2013. I searched the database for segment names that contained the term “Defense,” “Military,” “Federal,” “Government,” or “Govt.” This search process identified 39 unique parent companies. Each of these companies believes that their Federal/Government/Military segment is sufficiently different and important to warrant separate presentation in their financial statements.

---

6 Accounting Tools, http://www.accountingtools.com/questions-and-answers/what-is-segment-reporting.html, provides an excellent description of segment reporting on its website:

Under Generally Accepted Accounting Principles (GAAP), an operating segment engages in business activities from which it may earn revenues and incur expenses, has discrete financial information available, and whose results are regularly reviewed by the entity’s chief operating decision maker for performance assessment and resource allocation decisions. Follow these rules to determine which segments need to be reported:

- Aggregate the results of two or more segments if they have similar products, services, processes, customers, distribution methods, and regulatory environments.
- Report a segment if it has at least 10% of the revenues, 10% of the profit or loss, or 10% of the combined assets of the entity.
Table 1. The Sample Creation

Panel A. Federal Focus Firms

<table>
<thead>
<tr>
<th>Description</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active and Inactive Observations From Compustat Annual Updates—Segments</td>
<td>166,898</td>
</tr>
<tr>
<td>(Non-historical) 2010–2013</td>
<td></td>
</tr>
<tr>
<td>Parent Firms with Federal, Government, or Military Segment</td>
<td>39</td>
</tr>
<tr>
<td>Observations for Federal focus from Compustat, Fundamental Annual, 1993–2012</td>
<td>517</td>
</tr>
<tr>
<td>Less:</td>
<td></td>
</tr>
<tr>
<td>Outliers</td>
<td>473</td>
</tr>
<tr>
<td>Influential Observations</td>
<td>269</td>
</tr>
<tr>
<td>Total Observations</td>
<td>269</td>
</tr>
<tr>
<td>Number of Unique Companies</td>
<td>35</td>
</tr>
</tbody>
</table>

Panel B. Control Firms

<table>
<thead>
<tr>
<th>Description</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active and Inactive Observations From Compustat, Fundamental Annual 1993–2012</td>
<td>211,513</td>
</tr>
<tr>
<td>Less observations excluded because they are:</td>
<td></td>
</tr>
<tr>
<td>Foreign Companies</td>
<td>191,744</td>
</tr>
<tr>
<td>Missing or Infeasible data</td>
<td>89,214</td>
</tr>
<tr>
<td>Sales decr and SGA/COGS incr, or SGA &gt; Sales</td>
<td>80,096</td>
</tr>
<tr>
<td>Not in Fed Focus Industry or Year</td>
<td>49,741</td>
</tr>
<tr>
<td>Outliers</td>
<td>43,103</td>
</tr>
<tr>
<td>Influential Observations</td>
<td>39,539</td>
</tr>
<tr>
<td>Total Observations</td>
<td>39,539</td>
</tr>
<tr>
<td>Number of Unique Companies</td>
<td>6,991</td>
</tr>
</tbody>
</table>
I used the parent company identification to draw data from Compustat, Fundamental Annual for the years 1993–2012. At this stage, there were 517 observations. Later analysis trimmed outliers and removed influential observations. The final list of federal focus firms contains 269 observations for 35 unique companies.

My list of 35 federal focus firms and their associated industry, their Naics Sector, is provided in Table 2. While the majority of the firms are in manufacturing industries, there is a wide divergence of other industries represented, from construction to educational services.

Table 2. Federal Focus Companies by Naics Sector

<table>
<thead>
<tr>
<th>CONSTRUCTION-23</th>
<th>WHOLESALE TRADE - 42</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLUOR CORP*</td>
<td>AAR CORP*</td>
</tr>
<tr>
<td>KBR INC</td>
<td>NASH FINCH CO</td>
</tr>
<tr>
<td>MANUFACTURING -32, 33</td>
<td>INFORMATION - 51</td>
</tr>
<tr>
<td>USEC INC</td>
<td>SAPIENT CORP</td>
</tr>
<tr>
<td>ALLIANT TECHSYSTEMS INC*</td>
<td>FINANCE AND INSURANCE - 52</td>
</tr>
<tr>
<td>ANAREN INC</td>
<td>HEALTH NET INC*</td>
</tr>
<tr>
<td>BOEING CO*</td>
<td>PROFESSIONAL, SCIENTIFIC AND TECHNICAL - 54</td>
</tr>
<tr>
<td>CUBIC CORP</td>
<td>BAKER (MICHAEL) CORP*</td>
</tr>
<tr>
<td>EMS TECHNOLOGIES INC</td>
<td>CH2M HILL COS LTD</td>
</tr>
<tr>
<td>FEDERAL SIGNAL CORP</td>
<td>INTEGRAL SYSTEMS INC*</td>
</tr>
<tr>
<td>FLIR SYSTEMS INC</td>
<td>SAIC INC</td>
</tr>
<tr>
<td>GENCORP INC</td>
<td>URS CORP*</td>
</tr>
<tr>
<td>HARRIS CORP*</td>
<td>ADMINISTRATIVE AND SUPPORT - 56</td>
</tr>
<tr>
<td>II-VI INC</td>
<td>ENERGYSOLUTIONS INC</td>
</tr>
<tr>
<td>IROBOT CORP</td>
<td>KFORCE INC</td>
</tr>
<tr>
<td>ITT CORP*</td>
<td>EDUCATIONAL SERVICES - 61</td>
</tr>
<tr>
<td>MOOG INC -CL A</td>
<td>GP STRATEGIES CORP</td>
</tr>
<tr>
<td>NATIONAL PRESTO INDS INC</td>
<td></td>
</tr>
<tr>
<td>RAYTHEON CO*</td>
<td></td>
</tr>
<tr>
<td>SPARTON CORP</td>
<td></td>
</tr>
<tr>
<td>SYMMETRICOM INC</td>
<td></td>
</tr>
<tr>
<td>TEL-INSTRUMENT ELECTRONICS</td>
<td></td>
</tr>
<tr>
<td>TELEDYNE TECHNOLOGIES INC*</td>
<td></td>
</tr>
<tr>
<td>TRIMAS CORP</td>
<td></td>
</tr>
</tbody>
</table>

* A top 500 recipient of defense contracts for 2008.

As was mentioned previously, my approach to identifying firms with a federal focus differs from prior work. For instance, Wang and San Miguel (2012) used Fedspending.org to identify 112 publicly traded companies in the 500 companies with the largest dollar values of defense contracts awarded in 2008. Comparing their list to Table 2 shows that there are 12

---

7 Following ABJ, I use 20 years of data to estimate my models.
firms in common. Untabulated results show that the 12 common firms are substantially larger than the other 23 firms.\footnote{The 12 firms on both lists have average assets (sales) of 7,786 (8,927), while the 23 other firms in my sample have average assets (sales) of 995 (1,434)}

\textbf{Control Sample}

My control sample provides a benchmark to judge the performance of the federal focus firms. Table 1, Panel B provides the details.

The control sample began with the entire set of active and inactive Compustat firms for 1993–2012. I deleted foreign firms. I then removed observations missing data or with infeasible data (e.g., negative revenues). Next, I dropped observations if Sales decreased, but SGA or COGS rose. Following ABJ, I also removed observations if the firm’s SGA costs were greater than Revenues. I then deleted all year * industry observations with no federal focus observation in that year * industry.

Prior work has consistently found that sticky costs only show up in the estimation results after extensive data cleaning. For instance, ABJ trimmed 1\% of all variables and threw out all influential regression observations in order to generate results. My sample contains a later time period than ABJ and includes the Great Recession. It therefore contains many more outliers than ABJ. In order to obtain baseline results with SGA sticky costs, I removed the top and bottom 5\% of the log\left(\frac{\text{Revenue}_{it}}{\text{Revenue}_{it-1}}\right) observations and the top (bottom) 0.5\% of the DecrDum\_it * log\left(\frac{\text{Revenue}_{it}}{\text{Revenue}_{it-1}}\right) observations. Although I trimmed many observations, the range of variation in my retained data is comparable to prior work.

My final data reduction removed all observations that were influential in the SGA Sticky Cost regression.\footnote{Deleting observations for each regression separately generates qualitatively identical results.} My final set of controls consists of 39,539 observations over 20 years for 6,991 firms in nine industries.

My estimation sample combines the control firms with the federal focus firms. One important observation is that there are comparatively few federal focus (269) to control observations (39,539). A major concern is that the signal from the small number of federal focus observations could be drowned out by the large number of controls. Although this is a valid issue, the regression results generate statistically significant, and intuitive, estimates.

The imbalance between the number of control and federal focus observations may explain the knife-edge nature of my results. The federal focus signal is strong only under a tightly controlled set of data conditions.

\textbf{Descriptive Statistics}

Table 3, Panel A provides the descriptive statistics for my sample. Supplemental information shows that average total assets are comparable for federal focus and control firms (3,972 million versus 3,892 million), but that on average, federal focus firms have more employees (16,803 versus 5,556). In addition, federal focus firms have higher revenues, higher SGA expense, and greater COGS. These disparities are excellent reasons why equations use ratios to control for scale effects.
Table 3, Panel B runs the Wilcoxon non-parametric test for differences between variables in the control and federal focus sub-samples. The tests fail to reject that the dependent variable, $\log \left( \frac{S_{GA_{it}}}{S_{GA_{it-1}}} \right)$, and one of the independent variables, $\log \left( \frac{Revenue_{it}}{Revenue_{it-1}} \right)$, are drawn from the same distribution. However, the interaction term, $\text{DecrDum}_{it} \times \log \left( \frac{Revenue_{it}}{Revenue_{it-1}} \right)$, does show a significant difference between federal focus and control observations. This difference has two sources. First, federal focus firms have fewer revenue decreases than controls (11.9% versus 19.5%). Second, untabulated results show that, conditional on revenues falling, the federal focus firms have a smaller reduction in the log revenue ratio (-0.0787 versus -0.190). When revenues fall, federal focus firms are not hit as hard as the control firms.

Table 3. Descriptive Statistics and Pairwise Tests

Panel A. Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Combined N = 39,808</th>
<th>Federal Focus N = 269</th>
<th>Controls N = 39,539</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev</td>
<td>Mean</td>
</tr>
<tr>
<td><strong>Supplemental Information</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Assets(^1)</td>
<td>3,802.2</td>
<td>36.221</td>
<td>3,971.5</td>
</tr>
<tr>
<td>Number Employees(^2)</td>
<td>5.6322</td>
<td>21.460</td>
<td>16.030</td>
</tr>
<tr>
<td><strong>Components</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SGA(^1)</td>
<td>332.09</td>
<td>1,484.2</td>
<td>508.75</td>
</tr>
<tr>
<td>COGS(^1)</td>
<td>1,142.9</td>
<td>6,858.1</td>
<td>3,759.6</td>
</tr>
<tr>
<td>Revenue(^1)</td>
<td>1,794.1</td>
<td>9,296.1</td>
<td>4,749.2</td>
</tr>
<tr>
<td>DecrDum</td>
<td>0.1848</td>
<td>0.3861</td>
<td>0.1190</td>
</tr>
<tr>
<td><strong>Dependent Variable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(SGA(<em>{it}/SGA</em>{it-1})</td>
<td>0.0753</td>
<td>0.1814</td>
<td>0.0065</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(Revenue(<em>{it}/Revenue</em>{it-1})</td>
<td>0.1014</td>
<td>0.2064</td>
<td>0.0097</td>
</tr>
<tr>
<td>DecrDum *</td>
<td>-0.0369</td>
<td>0.1036</td>
<td>-0.0009</td>
</tr>
<tr>
<td>Log(Revenue(<em>{it}/Revenue</em>{it-1})</td>
<td>0.0007</td>
<td>0.0819</td>
<td>1</td>
</tr>
</tbody>
</table>

Panel B. Pairwise Tests for Dependent and Independent Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Federal Focus</th>
<th>Controls</th>
<th>F Value</th>
<th>Pr. &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(SGA(<em>{it}/SGA</em>{it-1})</td>
<td>0.0685</td>
<td>0.0754</td>
<td>0.3793</td>
<td>0.5380</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(Revenue(<em>{it}/Revenue</em>{it-1})</td>
<td>0.0897</td>
<td>0.1015</td>
<td>0.8674</td>
<td>0.3517</td>
</tr>
<tr>
<td>DecrDum *</td>
<td>-0.0009</td>
<td>-0.0371</td>
<td>19.124**</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

\(^1\) Millions  
\(^2\) Thousands

Now that I have described the data, I turn to the regression analysis.
Estimation Results

Table 4 contains the estimation results for the Selling, General, and Administrative (SGA) Sticky Cost model.\textsuperscript{10} The model uses a log-log specification, which means that all coefficients (except the constants) generate an elasticity.

Table 4. Sticky Cost Regression, Dependent Variable log[SGAt/SGA\textsubscript{(t-1)}]

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LRev\text{ratio} \ a_1$</td>
<td>0.6922*** (0.000)</td>
</tr>
<tr>
<td>$Decr^*LRev\text{ratio} \ a_2$</td>
<td>-0.0628*** (0.000)</td>
</tr>
<tr>
<td>$FSeg \ a_3$</td>
<td>-0.0113** (0.029)</td>
</tr>
<tr>
<td>$FSeg^*LRev\text{ratio} \ a_4$</td>
<td>0.1380*** (0.000)</td>
</tr>
<tr>
<td>$FSeg^* Decr^*LRev\text{ratio} \ a_5$</td>
<td>-0.3136** (0.000)</td>
</tr>
<tr>
<td>Adj $R^2$</td>
<td>0.5855</td>
</tr>
<tr>
<td>N</td>
<td>39,808</td>
</tr>
</tbody>
</table>

My control sample’s behavior is captured by coefficients $a_1$ and $a_2$. The control results demonstrate sticky SGA costs, though the costs are less sticky than in older samples.\textsuperscript{11} Combining the correct coefficients shows that a 1% increase in the revenue ratio ($a_1$) leads to a 0.692% increase in the control firms’ SGA, while a 1% decrease ($a_1+a_2$) leads to a 0.629% decrease in SGA. The SGA costs are 0.063% sticky ($a_2$). I can benchmark the magnitude of these effects by evaluating the elasticity at the mean sample values. When revenues (average 1,794 million) increase/decrease by 1%, they change by 17.94 million. When revenues rise (fall) by 17.94 million, SGA costs (average 332.1 million) increase (fall) by 2.30 (2.09) million. At the median combined sample values, SGA costs stick by 0.21 million, or roughly $210,000.

\textsuperscript{10} The model is estimated with fixed industry effects and uses robust standard errors. The VIF scores show no significant multicollinearity.

\textsuperscript{11} Anderson, Banker, and Janakiraman (2003) estimate cost stickiness for 20 years, from 1979 to 1998 for all but the financial services industry. They find (Table 2, Model [I]) a comparable increase in SGA costs when revenues rise (0.5459 versus my 0.6922), but a larger decline when costs fall (-0.1914 to my -0.0628). I conjecture the differences are due to my restriction to a subset of industries as well as the different time periods.
The impact of a federal focus on the SGA costs is identified through the interaction terms. Combining the correct coefficients for federal-focused companies \((a_1+a_3)\), when the revenue ratio rises by 1%, the SGA ratio rises by 0.828%, which is 0.136% more than for the controls \((a_4)\). When revenues rise, federal focus companies have a \(0.136/.692 = 19.7\%\) increase in SGA over the controls. This incremental rise could reflect greater paperwork requirements/fulfillment costs for federal focus firms. Combining the correct coefficients \((a_1+a_2+a_3+a_4)\), when revenues fall by 1%, the SGA ratio for federal focus falls by 0.447%, which is 0.182% less than for the controls \((a_1+a_2)\). When revenues fall, federal focus companies have a \(0.182/.629 = 28.9\%\) smaller decrease in SGA costs than control companies. The slower fall could reflect greater fixed costs for federal focus companies, in particular, greater fixed staff costs in the procurement process.

The dollar impact for federal focus companies can be evaluated using the mean sample values. When revenues rise by 1% (17.94 million), SGA costs rise by 2.75 million for federal focus companies versus 2.30 million for the controls—a 0.45 million cost difference. When revenues fall by 1% (17.94 million), SGA costs fall by 1.48 million for federal focus companies, and by 2.09 million for the controls—a 0.61 million cost difference. The dollar value of the federal focus difference is understated since federal focus firms tend to be much larger than the control firms.\(^{12}\)

Figure 1 illustrates the qualitative behavior of SGA costs for both the controls and the federal focus companies. If a firm is federal focused, then SGA costs rise faster when revenue increases, but fall slower when revenue decreases.

\(^{12}\) If the numbers are evaluated at the median federal focus firm value (SGA 508.75), then a 1% revenue increase would cause SGA for controls to rise by 3.522 and federal focus by 4.213, a difference of 0.691 million. A 1% revenue decrease would cause SGA for controls to fall by 3.202 million, and federal focus by 2.273 million, a difference of 0.929 million.
**Figure 1. The Change in the SGA Ratio as a Result of Revenue Increases and Decreases**

**Robustness Checks**

My robustness tests suggest that my results are delicate, are knife-edge. Reducing the time period from 20 to 15 (10) years generates qualitatively similar control coefficients, but the federal focus coefficients become insignificant. Alternatively, using the median industry values as controls leads to all coefficients becoming insignificant.

**Conclusions**

Federal government contractors are qualitatively different than other firms (Kovacic, 1992). Contractors complain about excessive paperwork requirements, fixed margins, and long lead times to obtain contracts. I examine whether these problems are large enough to show up in the published financial accounting data. I use the behavior of SGA costs on firms’ published Income Statements to address this issue.

An important innovation in my paper is the creation of a federal focus sample.

I use Compustat’s Segment database to identify firms that have a business segment labeled *federal, government, or military*. These federal focus companies believe that their federal government business is so important that they have built their organization structure around it. While there is some overlap between my list of federal focus companies and the firms with the greatest sales to the federal government, roughly two thirds of my federal
focus firms are not on the list of the largest 500 federal dollar contractors. My sample contains a different, and possibly superior, set of companies than offered by prior work.

Once I generated the federal focus sample, I estimated an SGA sticky cost regression where I included interaction terms for federal focus firms. The results show that the controls have mild SGA cost stickiness. Raising revenue 1% leads to a 0.70% rise in SGA costs, while lowering revenues 1% leads to a 0.63% decline in SGA costs. My main results show that there are significant differences for federal-focused firms. Raising revenues 1% leads to a 0.82% rise in SGA costs, a 20% increase over the controls. When revenues rise, federal-focused firms have a higher ramp up in SGA costs, possibly due to increased paperwork for fulfilling government orders. In the same vein, lowering revenues 1% leads to a 0.45% drop in SGA costs, about 29% below the reduction for the control group. SGA costs for federal focus companies are much stickier than for the controls. The reluctance to reduce SGA costs is consistent with federal focus firms maintaining their core federal procurement process in a downturn.

Robustness tests suggest that my sample is delicate. Changing the time frame or the approach to generating controls leads to no significant results. This delicacy may be due to the small number of federal focus observations relative to control observations.

My paper provides evidence that SGA costs behave differently for government contractors. Balakrishnan, Petersen, and Soderstrom (2004) provided evidence that the magnitude of the change matters. They showed that very large changes in costs lead to greater responsiveness (less stickiness) than smaller changes, and argued that transaction costs will dampen or remove small changes, but not affect larger changes. While similar behavior might hold for federal contractors, I cannot investigate this issue. My sample contains too few observations to analyze this issue.

My analysis looks at the behavior of firms which have a federal focus. However, it does not investigate one other important aspect of government contracting: the form of the contracts. A large body of theoretical literature explores the relationship between optimal contracts and information in procurement (Laffont & Tirole, 1986; Rogerson, 1994). For instance, cost plus contracts may lead to improved risk sharing, but can lead firms to shift costs from commercial to government contracts (Chen & Gunny, 2014; Rogerson, 1992). In contrast, fixed price contracts provide no incentives to shift costs, but may lead firms to underinvest in fixed assets to support the contract. My SGA sticky cost results are consistent with the cost shifting of cost plus contract, but I have no information as to the actual contract form. Future work could connect the contract form with my analysis.

References


**Acknowledgments**

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.

Thursday, May 14, 2015

| 1:45 p.m. – 3:15 p.m. | Chair: **Dr. John Burrow**, Deputy Assistant Secretary of the Navy for Research, Development, Test, & Evaluation

**Addressing the Barriers to Agile Development in the Department of Defense: Program Structure, Requirements, and Contracting**

Su Chang, MITRE Corporation  
Pete Modigliani, MITRE Corporation

**Finding the “RITE” Acquisition Environment for Navy C2 Software**

George Galdorisi, SPAWAR Systems Center Pacific  
Amanda George, SPAWAR Systems Center Pacific  
Michael Morris, SPAWAR Systems Center Pacific

**Use of Automated Testing to Facilitate Affordable Design of Military Systems**

Valdis Berzins, NPS  
Paul Van Benthem, SPAWAR  
Christopher Johnson, SPAWAR  
Brian Womble, DASN (RDT&E)

---

**Dr. John Burrow**—serves as Deputy Assistant Secretary of the Navy for Research, Development, Test and Evaluation (DASN RDT&E) under the Assistant Secretary of the Navy for Research, Development and Acquisition (ASN (RD&A)). Dr. Burrow is responsible for executive oversight of all matters related to RDT&E Budget Activities, Science and Engineering, Advanced Research and Development, Prototyping and Experimentation, and Test and Evaluation. He is also responsible for oversight and stewardship of the Department of Navy Research and Development Establishment including naval laboratories, warfare centers and systems centers.

Dr. Burrow has more than 30 years of Federal Service as an acquisition professional, systems engineer and technical leader. He has held numerous senior leadership positions including Executive Director, Marine Corps Systems Command (MCSC); Deputy Commander (MCSC), Systems Engineering, Interoperability, Architectures and Technology; Head, Naval Surface Warfare Center, Force Warfare Systems Department; Systems Engineering Director, Program Executive Office, Integrated Warfare Systems; Director, Marine Corps Amphibious Combat Vehicle Assessment Team; and Director, Navy Small Surface Combatant Task Force.

Dr. Burrow was appointed to the Senior Executive Service in December 2004. He is a Certified Level III Acquisition Professional in the Advanced Systems Planning, Research, Development and Engineering (SPRDE) and Program Management (PM) acquisition career fields. Dr. Burrow holds a Bachelor of Science in Mathematics from the University of Mississippi (1983), a Master of Public Administration from Virginia Polytechnic Institute and State University (1997), and a Doctorate of Management from the University of Maryland University College (2009).
Addressing the Barriers to Agile Development in the Department of Defense: Program Structure, Requirements, and Contracting

Su Chang—is a principal economic and business analyst at The MITRE Corporation, specializing in IT acquisition. As MITRE’s Contracting Technical Team co-leader, she fosters collaboration and professional development of MITRE’s contracting and acquisition analysts. Prior to joining MITRE, Chang was a senior contract negotiator with the Missile Defense Agency, and is a graduate of the Department of Interior, Government-wide Acquisition Intern Program. She holds a BS in economics from the University of Utah, and an MA in U.S. foreign policy from American University. She is DAWIA Certified Level III in contracting.

Pete Modigliani—is the acquisition innovation area lead at The MITRE Corporation. He supports DoD acquisition and CIO executives’ strategic initiatives in Agile, cyber, IT, and services acquisition. He manages a research portfolio to foster innovative acquisition solutions. Previously, as an assistant vice president with Alion, he supported the Air Force Acquisition Executive on C4ISR systems. As an Air Force program manager, he developed strategies for billion-dollar acquisitions. Pete holds a BS in industrial engineering from the Rochester Institute of Technology and an MBA from Boston College. He is DAWIA Certified Level III in program management. [pmodigliani@mitre.org]

Abstract

Program managers and executives in the Department of Defense (DoD) have struggled for years to tailor the acquisition framework to promote delivery of information technology (IT) capabilities in small, frequent releases—the approach that characterizes Agile development. DoD acquisition professionals increasingly recognize the potential of Agile methods, but do not know how to apply Agile within the unique and complex defense acquisition environment. Several aspects of the defense acquisition process have proven especially challenging in the implementation of Agile practices. For example, the lack of knowledge about how to tailor the DoD Instruction (DoDI) 5000.02 process for an Agile development can deter a program from considering the use of Agile techniques in the first place. Many DoD IT acquisition programs are unfamiliar with the IT Box requirements concept, and thus cannot take advantage of its flexibilities to enable Agile development. In addition, long contracting timelines and the tendency to lock down Agile requirements in a contract have become barriers to implementing the speed and flexibility necessary for successful Agile adoption. This paper offers specific acquisition solutions and strategies to address these identified “high barriers” to Agile development in DoD.

Introduction

Agile software development practices integrate planning, design, development, and testing into an iterative lifecycle to deliver software at frequent intervals. Structuring programs and processes around small, frequent Agile releases enables responsiveness to changes in operations, technologies, and budgets. These frequent iterations effectively

1 Approved for public release; distribution unlimited. Case Number 15-0905. The author’s affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE’s concurrence with, or support for, the positions, opinions or viewpoints expressed by the author.
measure progress, reduce technical and programmatic risk, and respond to feedback and changes more quickly than traditional waterfall methods (Modigliani & Chang, 2014).

While the commercial sector has broadly adopted Agile development to rapidly and dynamically deliver software capability, Agile has just begun to take root in DoD acquisitions (Lapham et al., 2010; Northern et al., 2010). A dozen or more DoD IT acquisition programs have incorporated Agile concepts and practices. These early adopters, like any new venture, have experienced mixed results. Furthermore, despite the early adoption of Agile across several DoD IT acquisition programs, the DoD has issued no formal guidance and training on DoD Agile practices. Many acquisition professionals see the value and promise of Agile, yet struggle to incorporate it effectively in the Defense Acquisition Framework. Given that Agile in many ways differs so radically from the DoD’s traditional development practices, programs interested in using Agile encounter several challenges and barriers within the DoD acquisition system (Broadus, 2013).

MITRE performed initial research to examine the leading Agile methodologies and commercial practices and explore how DoD acquisition structure and processes could be tailored to adopt Agile. The resulting Defense Agile Acquisition Guide provides acquisition professionals guidance and instruction for Agile adoption. Following the release of the guide in February 2014, MITRE conducted further research to capture best practices and lessons learned from early adopters across the DoD and other federal agencies. The research, based on years of experience and collaboration across Agile and IT acquisition communities, refined and extended strategies for tailoring each functional area of acquisition. This paper focuses specifically on three of the most difficult barriers to successful DoD Agile adoption: program structure, requirements, and contracting. The DoD can address these barriers by utilizing a proactively tailored Agile acquisition model, implementing an IT Box requirements process, and utilizing the flexible contracting approaches described in this paper.

The first half of this paper provides an overview of the Agile development process and identifies some of the primary challenges in adapting commercial Agile practices for DoD implementations. Next, the paper examines prerequisites for effective adoption of Agile practices in the DoD. The remaining sections describe each of the three “high barrier” problem areas and offer specific recommendations that the DoD can use today to overcome these challenges.

Background

Agile Development Overview

Agile software development emerged in 2001 after 17 industry leaders created the Agile Manifesto to design and share better ways to develop software (Agile Manifesto, n.d.). Agile prioritizes early and continuous deliveries of working software; adapts easily to changing requirements; depends on small, empowered teams; and promotes active user involvement during development.

Agile development can be distilled into four core elements:

2 A copy of the guidebook can be obtained at http://www.mitre.org/publications/technical-papers/defense-agile-acquisition-guide-tailoring-dod-it-acquisition-program
• Focusing on small, frequent capability releases
• Valuing working software over comprehensive documentation
• Responding rapidly to changes in operations, technology, and budgets
• Actively involving users throughout the development process to ensure high operational value

The foundation of Agile is a culture of small, dynamic, empowered teams actively collaborating with stakeholders throughout product development. Agile development requires team members to follow disciplined processes that require training, guidance, and openness to change (GAO, 2012). While Agile does impose some rigor, the method does not consist of simply following a set of prescribed processes, but instead allows dynamic, tailored, and rapidly evolving approaches that suit each organization’s IT environment.

Various Agile methods (e.g., Scrum, Extreme Programming (XP), Kanban, Test Driven Development) have emerged, each with unique processes, terms, and techniques (Modigliani & Chang, 2014). These methods focus on the development team and associated stakeholders. Agile Acquisition extends Agile development practices beyond the contractor development team to the government acquirer, users, and other stakeholders. It requires that both agencies and contractors change many acquisition roles, processes, and culture, thereby fostering a close government–industry partnership (Balter, 2011; Lapham et al., 2010). This, in turn, demands investments in time, training, and continuous improvement to pay long-term dividends. While both practical experience and research findings strongly indicate the value of Agile acquisition for many IT development programs, this approach may not be appropriate in all cases (Lapham et al., 2010). Programs should consider Agile Acquisition when

• Users can decompose requirements into small tasks for iterative development.
• The operational environment can support small, frequent capability deliveries.
• Users can engage throughout development to capture concepts of operations (CONOPS) and provide feedback on demonstrated capabilities.
• Program can use existing infrastructure and focus development on the application layer.
• Industry partners are available with relevant domain expertise in Agile practices.
• Milestone Decision Authority supports Agile development practices and tailored processes.

**Agile and the DoD Acquisition Environment**

Despite the success that Agile development has achieved in the private sector, commercial implementation of Agile does not directly translate to Agile adoption in the federal sector. The barriers to program structure, requirements, and contracting often stem from these key differences. First, the government must adhere to a set of rigorous policies, statutes, and regulations that do not apply to the same degree to the commercial sector (Lapham et al., 2011). Following the rules that govern federal acquisition often involves a bureaucratic, laborious, and slow process that greatly influences how effectively the DoD can implement Agile. Second, the commercial sector has a different stakeholder management process than the government. Private firms are accountable to an internal and layered management structure that usually goes no higher than a corporate board of directors; the few possible external stakeholders (e.g., labor unions) rarely cause frequent
and major disruptions. The government bureaucracy has layers upon layers of stakeholders with a high degree of influence that can create frequent and significant disruptions. Everything from a change in the political administration to budget sequestration can exert significant external influence on a DoD program. Lastly, the bureaucratic layers of government make it difficult to empower Agile teams to the same extent as in the private sector. The commercial sector has considerable latitude to make adjustments throughout the course of the development because companies closely link accountability, authority, and responsibilities to push decision-making to the lowest levels. The government’s tiered management chain of command makes it difficult for the Agile team to make decisions quickly and unilaterally.

The above comparisons demonstrate the need for the DoD to tailor Agile processes to its unique set of policies and laws. Herein lies the fundamental issue with Agile adoption in the DoD. The practices, processes, and culture that have made Agile development successful in the commercial sector often run counter to the current practices, processes, and culture in the long-established defense acquisition enterprise (Broadus, 2013). In many ways, the acquisition environment needed to execute Agile development is the opposite of the acquisition environment in place today.

- The small, frequent capability releases that characterize the Agile development approach directly contrast with the traditional DoD acquisition model designed for a single big-bang waterfall approach (Broadus, 2013). Currently, every step in the acquisition system must be extensively documented and approved prior to execution. For example, according to DoDI 5000.02, a DoD IT acquisition program must meet 34 statutory and regulatory documentation requirements prior to entering Milestone A (Defense Acquisition University, 2015), whereas Agile emphasizes working software over comprehensive documentation (Lapham, 2012).

- Agile also enables rapid response to changes in operations, technology, and budgets. By contrast, the DoD requires budgets, requirements, and acquisitions to be planned up front, sometimes several years in advance of execution, and changing requirements, budgets, and strategies during the execution process is disruptive, time-consuming, and costly (Modigliani & Chang, 2014).

- Lastly, Agile values active involvement of users throughout the development process to ensure high operational value, and continuously re-prioritizes the ongoing requirements process on the basis of feedback from the user community on deployed capabilities. Today’s DoD requirements process is static, rigid, and limits active user involvement and feedback during the development process (Lapham et al., 2010).

Given these key differences, the DoD has been ill prepared to adopt Agile development practices and in fact Agile implementations so far have not always succeeded. Some early DoD adopters attempted what they thought or promoted as “Agile,” yet they did not incorporate some of the foundational Agile elements into their structures or strategies. This resulted partly from the lack of definition and standardized processes for Agile in the federal sector. In some cases, programs implemented a few Agile principles, such as
breaking large requirements into smaller increments, but did not integrate users during the development process to provide insight or feedback. Other programs structured capability releases in a time-boxed manner, yet did not understand what to do when releases could not be completed in time.

Adopting only a handful of Agile practices without a broader Agile strategy often fails to achieve desired results (GAO, 2012). For example, one DoD early adopter initially attempted to implement Agile practices by breaking large requirements into several four-week sprint cycles. However, the program lacked high-level agreement on what to develop in each cycle, and did not have a robust requirements identification and planning process in place. Furthermore, the program lacked an organized user community and active user-participation throughout the development process—a fundamental Agile tenet. As a result, the Agile processes quickly degenerated and the program only delivered 10% of its objective capability after two years of failed Agile development attempts. The program finally retreated to a waterfall-based process. It simply could not execute the Agile strategy without the proper environment, foundation, and processes in place. On the other hand, the DoD has recorded some significant successes with Agile, such as the Global Combat Support System–Joint (GCSS-J) program, which has routinely developed, tested, and fielded new functionality and enhanced capabilities in six-month deployments (Defense Information Systems Agency, 2015).

**Prerequisites for Agile Adoption**

The Agile model represents such a radical change in the way the DoD conducts business that the DoD must actively rethink how programs are managed and structured to support Agile (Modigliani & Chang, 2014). This requires restructuring the current acquisition environment (i.e., policies, processes, and culture) to enable success.

As a starting point, the DoD should adopt a common understanding of Agile and identify the underlying set of values that describe the purpose and meaning of DoD Agile practices. The authors propose the following guiding principles for DoD Agile adoption:

1. **Focus on small, frequent capability releases to users**—Smaller releases are easier to plan, present lower risks, and are more responsive to changes. Projects should focus on delivering working software as the primary objective.

2. **Embrace change**—Projects must allow for changes to scope and requirements based on operational priorities, user feedback, early developments, budgets, technologies, etc. This requires flexible contracts, strong collaboration, and rigorous processes. Projects should plan early and then adapt based on current conditions.

3. **Establish a partnership between the requirements, acquisition, and contractor communities**—Projects should foster active collaboration on operations, technologies, costs, designs, and solutions. This requires committed users who contribute to development, tradeoff discussions, and regular demonstrations of capabilities.

---

3 A time-box is a fixed time period allocated to each planned activity. For example, within Agile, a sprint is often time-boxed to a 4–6 week time period or a release is time-boxed for a 4–6 month time frame.
4. **Rely on small, empowered, high-performing teams to achieve great results**—Organizing around each release with streamlined processes and decisions enables faster deliveries that are more successful.

5. **Leverage a portfolio structure**—Individual programs and releases can deliver capabilities faster by using portfolio or enterprise strategies, processes, architectures, resources, and contracts.

These tenets align with the recommended set of principles in the Government Accountability Office (GAO) report on *Effective Practices and Federal Challenges in Applying Agile Methods*. They center on the Agile Manifesto themes of small, frequent capability releases, a dynamic requirements process that allows for the continuous prioritization of requirements, active involvement from the user community throughout the development process, and commitment to delivering working software based on a time-boxed schedule. Some efforts may succeed in implementing only a subset of these themes and delivering effective software solutions; however, one could argue that this would not constitute a pure Agile development.

The DoD would benefit from defining and standardizing Agile-based practices to ensure a Department-wide consistent and common understanding of what constitutes an Agile-based DoD program or project (Lapham et al., 2010). Today, many efforts are inaccurately labeled as Agile, leading to misunderstanding and misrepresentation of Agile principles. After defining the principles, the DoD needs to provide detailed guidance to the acquisition community that describes how to execute the Agile acquisition processes within DoD acquisition regulations and laws (Broadus, 2013). This level of detailed process-level guidance falls outside the scope of this paper, but the *Defense Agile Acquisition Guide* offers further details on the guidance needed for the DoD to make Agile adoption effective and widespread. This paper centers on the aspects of the DoD acquisition process that have proven most problematic when implementing Agile development concepts. The following sections focus on three of the most difficult barriers for DoD Agile adoptions: program structures, requirements, and contracting.

### Structuring an Agile Program

Structuring a program for Agile development differs significantly from structuring an IT program around a traditional development methodology. Traditional waterfall programs usually have discrete acquisition phases driven by milestone events to deliver a large capability. Agile is more dynamic and requires the program to be structured to support multiple, small capability releases (Lapham, 2012).

Structuring an Agile program in this way represents a fundamental first step in developing a strategy for program-level Agile adoption. This activity requires the program to make significant adaptations to the traditional DoDI 5000.02 program structures and acquisition processes to support Agile development timelines and objectives (Modigliani & Chang, 2014). Although the DoDI 5000.02 acquisition policy places heavy emphasis on tailoring acquisition models to meet program needs, programs often do not know how to do so effectively and receive approval from process owners (Modigliani & Chang, 2014). It takes years of experience to truly understand the nuances involved in tailoring an acquisition program. Given the radical differences between Agile and a traditional development model, programs often view this activity as too complicated and therefore fail even to consider the Agile development process for a program. Programs must be designed in such a way that they not only meet all the DoDI 5000.02 statutory and regulatory requirements, but are also executable and marketable to senior acquisition executives who may be unfamiliar with the details of the Agile process. The following sections describe a recommended approach to
structuring an Agile DoD program, starting with the process to structure an Agile release and building on this concept to develop a fully tailored DoD Agile acquisition program.

**Agile Releases and Potential Program Structures**

When developing an Agile program structure, a program should first decide how to structure its releases. The release represents the core element of the program structure, guiding how frequently the program delivers capabilities to the warfighters. The length of each release depends upon operational, acquisition, and technical factors. As a general guideline, most releases should take less than 18 months, with a goal of 6–12 month timelines. Program offices should tailor acquisition processes to support these release timelines. In some cases, this requires redesigning key acquisition processes around a 6–12 month release rather than a 5–10 year increment.

Each release comprises multiple sprints and a final segment for release testing and certification. Each sprint, in turn, includes design, development, integration, and test, and culminates in demonstration of capabilities to users and other stakeholders. Developers may be required to deliver interim code to the government at the end of each sprint or multiple sprints. The government can integrate the interim code into its software environment for testing and operational assessments. Figure 1 shows a potential 6-month release structure with five monthly sprints.

![Figure 1. Potential Six-Month Release Structure](image)

Programs must adjust the length of the sprints and releases as conditions warrant. The key is to establish consistent, time-boxed releases, ideally small and frequent to allow for iterative development that responds easily to changes.

![Figure 2. Potential 12-Month Release Structure](image)

After determining a release strategy, each effort should tailor its programmatic and acquisition processes to effectively enable Agile development practices. Figure 3 illustrates one potential structure at a top level. In this approach, requirements, technology, and architecture development are continual processes rather than sequential steps in early acquisition phases. Each release involves a series of sprints to iteratively develop and test a capability, ultimately leading to capability deliveries to the warfighter every six months upon approval. Instead of bounding development via a series of increments with Milestones B and C at each end, development thus becomes a continual process. Semi-annual reviews with
senior leadership and other key stakeholders ensure transparency into the program’s progress, plans, and issues. Programs provide additional insight to executives via monthly or quarterly reports and other status meetings.

### Figure 3. Potential Agile Structure

A core theme throughout DoDI 5000.02 is the tailoring of program structures and acquisition processes to meet the needs of the individual program. The policy includes several acquisition models to consider, such as Model 2 for defense-unique software, Model 3 for incrementally fielded software, and hybrid Model B for software dominant programs. Figure 4 shows a proactively tailored acquisition model based on the three software models in DoDI 5000.02. The *Defense Agile Acquisition Guide* contains more detail about the structure and accompanying acquisition processes for DoD Agile adoption.
Agile Requirements Process

The Agile requirements process values flexibility and the ability to reprioritize requirements as a continuous activity based on user inputs and lessons learned during the development process. In contrast to current acquisition practices, the Agile methodology does not force programs to establish their full scope, requirements, and design at the start, but assumes that these will change over time.

At present, the Joint Capabilities Integration Development System (JCIDS) process guides the DoD requirements process. The traditional JCIDS process, based on lengthy and labor-intensive efforts to capture and define requirements, prevents agility (Lapham et al., 2011). The DoD has recognized that this process was particularly inappropriate to IT development because of the rapid pace of change in IT compared with the JCIDS requirements definition timeline. As a result, the DoD updated the JCIDS by approving an “IT Box” to better accommodate the dynamic nature of IT and the shortened timelines required to rapidly field IT-enabled operational capabilities (Office of the Secretary of Defense, 2010). The IT Box describes the operational performance and life-cycle affordability bounds of the program. The boundaries imposed by the “Box” expedite program initiation and streamline oversight by reducing return trips to the JROC for change approval.

However, even with the introduction of the IT Box model to provide more flexibility in the requirements process, many programs still struggle with how to apply this model to their IT development programs. As programs strive to structure their programs around Agile-based concepts as described in the previous section, they find it further confounding to figure out how to apply the IT Box model to satisfy the JCIDS requirements process. The following section contains specific recommendations on how to apply the IT Box concept to a DoD Agile development program.
Applying the IT Box Model for Agile Development

In the JCIDS IT Box model, an acquisition program develops an “IS-Initial Capabilities Document (ICD)” for JROC approval, while the traditional Capability Development Documents (CDDs) and Capability Production Documents (CPDs) are no longer required. Figure 5 illustrates the four sides of the IT Box identified in the IS-ICD.

As long as programs operate within these four sides of the IT Box, they need not return to the JROC for approval or oversight. In lieu of CDDs and CPDs, programs can develop Requirements Definition Packages (RDPs) to capture a subset of the IS ICD scope and/or Capability Drop (CD) documents for smaller items such as applications (see Figure 6). Most important, the requirements documents are designed for a smaller scope of work and approval at a lower level. This flexibility and streamlining of IT requirements enables Agile development within a DoD program. Programs should take advantage of this and avoid developing a CDD or CPD. Managers can formulate the requirements process for the overarching acquisition using the JCIDS IT Box process to build in flexibility from a high-level operational standpoint. Once an Agile approach has been designed into the program, programs must ensure they establish a flexible process for managing requirements from a functional capability standpoint (Modigliani & Chang, 2014).

Figure 5. IT Box
(Chairman of the Joint Chiefs of Staff [CJCS], 2012)

As long as programs operate within these four sides of the IT Box, they need not return to the JROC for approval or oversight. In lieu of CDDs and CPDs, programs can develop Requirements Definition Packages (RDPs) to capture a subset of the IS ICD scope and/or Capability Drop (CD) documents for smaller items such as applications (see Figure 6). Most important, the requirements documents are designed for a smaller scope of work and approval at a lower level. This flexibility and streamlining of IT requirements enables Agile development within a DoD program. Programs should take advantage of this and avoid developing a CDD or CPD. Managers can formulate the requirements process for the overarching acquisition using the JCIDS IT Box process to build in flexibility from a high-level operational standpoint. Once an Agile approach has been designed into the program, programs must ensure they establish a flexible process for managing requirements from a functional capability standpoint (Modigliani & Chang, 2014).

---

4 CDDs and CPDs are traditional JCIDS requirements documents that describe the program and program increment requirements.

5 Services and requirements oversight organizations have the flexibility to identify alternative names for these documents, along with their scope, content, and approval processes.
With the IT Box construct in place and the appropriate documentation requirements fulfilled, programs can manage the technical requirements in an Agile environment via program, release, and sprint backlogs. Backlogs could take the form of databases, Excel spreadsheets, or Agile development software tools. The product owner, the person responsible for requirements, actively manages (grooms) program and release backlogs, working with the user community and other stakeholders to identify the greatest level of detail for the highest priority requirements.

Figure 7 shows the relationships among the program, release, and sprint backlogs. The program backlog contains all desired functionality and requirements. A release backlog typically comprises the highest priority requirements from a program backlog that a team can complete within the established timeframe. A sprint then addresses the highest priority requirements from the release backlog. Once the development team commits to the scope of work for a sprint, that scope is locked. Sprint demonstrations conducted by the contractor at the end of a sprint may identify new features or defects that the team would add to the release or program backlogs.

---

6 A program backlog is the primary source of all requirements/desired functionality for the program. A release backlog is a subset of the program backlog listing features intended for the release. A sprint backlog is a subset of the release backlog listing the user stories to implement in the sprint.
The product owner, actively collaborating with users and stakeholders, is responsible for grooming the backlog to ensure the content and priorities remain current as teams receive feedback and learn more from developments and external factors. Users and development teams may add requirements to the program or release backlog or shift requirements between them. The release and development teams advise the product owner on the development impacts of these decisions, while users advise the release team about the operational priorities and impacts. To address a specific user story, the program must understand dependencies on existing or planned capabilities. Some programs may turn to a Change Control Board to make some of the larger backlog grooming decisions. The use of this model, combined with the IT Box structure, can help set a DoD Agile acquisition program on the right path for implementation (Modigliani & Chang, 2014).

Contracting for Agile Development

This section summarizes the difficulties of executing Agile development in the current government contracting environment and suggests available options.

Challenges

Contracting for Agile development has proven tremendously difficult not only for the DoD but also for many other federal agencies. The July 2012 GAO Report on Effective Practices and Federal Challenges in Applying Agile Methods cites “Procurement practices may not support Agile Projects” as a key challenge area. Contracting for Agile development presents a unique challenge to the government not often encountered in the private sector because commercial firms often rely on in-house staff to execute the Agile practices, whereas the government must obtain Agile development support through a contract arrangement.

This poses several challenges for the government. First, the government contracting process emphasizes competition and is guided by a set of policies and laws articulated in the Federal Acquisition Regulation (FAR). Government programs cannot simply choose any Agile development contractor they like, but must follow a specific set of contracting processes and protocols to obtain contracted support in a fair and transparent manner. These government contracting laws and regulations have resulted in long contracting timelines that in themselves pose significant difficulty for government implementation of Agile. A competitive IT contract can take over a year to award. This prevents execution of the Agile development process, which relies on short delivery cycles and time-boxed schedules (Lapham et al., 2011).

Next, the government contracting process requires programs to define the contract requirements upfront in a Statement of Work (SOW) or Performance Work Statement (PWS) so that a contractor can prepare a technical and cost proposal against the SOW/PWS requirements. The government uses the contractor’s proposal to determine the contract
scope, schedule, and cost. Herein lies one of the biggest obstacles to Agile implementation. One of the key tenets of Agile development is a dynamic requirements process that does not lock down requirements. The government therefore confronts the very difficult challenge of figuring out how to define requirements in a SOW/PWS to award the development contract, without locking down the technical requirements to a point where the contractor has no flexibility during the execution of the Agile development process (Balter, 2011).

Following contract award, successful Agile development depends on the ability to reprioritize requirements as program staff “learn” throughout the development process and re-scope the effort as needed. Today, however, post-award management of the contract is often inconsistent. In some cases, the contractor has minimal oversight and management and government–contractor interaction occurs only during infrequent reviews. By contrast, Agile requires very close management of the government–contractor relationship, with frequent, often daily, interaction between them.

Lastly, the award of a government contract today often relies on the strength of the proposed technical solution. Under Agile, the government and contractor together determine the technical solution in the course of executing the Agile development process. Thus, contract award should be based on the strength of the development team and the team’s experience using Agile practices.

Solutions

Given the disparity between traditional contracting practices and the needs of Agile, the government has encountered difficulties in contracting for Agile development. However, programs should consider the following solutions.

First, programs must plan contracts well in advance of the proposed Agile development. In many cases today, contracting can become the long-lead item in the development process if it is not properly considered in the upfront planning process.

Second, the program must determine if it will use a service or a product contract. A service contract is highly recommended because this vehicle would provide the program with greater flexibility to modify requirements along the development process (Modigliani & Chang, 2014). A service contract is more flexible for Agile efforts than a product contract because it describes requirements in terms of the people and time required to execute the development process rather than locking down the technical details of the end-product deliverable. However, this strategy assumes the program is the lead systems integrator and is responsible for overall product rollout and delivery. If the government expects the contractor to act as the systems integrator, determine the release schedule, and be accountable for overall product delivery, then a product-based contract in which the government describes overall delivery outcomes and objectives is more practical. However, this scenario would make it difficult for the government to execute a true Agile process, because changes to requirements in the course of development, or a change to the delivery schedule, will require a contract negotiation that could affect the Agile process. If the government does execute a product-based contract, it should pursue an indefinite-delivery, indefinite-quantity (IDIQ) contract vehicle and define product-based task orders based on either the release or the sprint level, depending on which level has the best-defined technical requirements (e.g., user stories). The program must carefully balance the advantages of a service versus a product contract based on a determination of government versus contractor responsibilities for the Agile processes.

Next, the program must determine the type of contract vehicle and strategy. Some cases require a separate stand-alone contract; in others, the government could leverage an existing contract vehicle. Programs must conduct thorough market research to determine if
an existing contract vehicle could meet their needs. When analyzing existing contract vehicles, a program must review the contract scope to ensure it can support Agile processes and evaluate the capabilities of the contract awardees to determine if they have Agile expertise and experience and if the labor categories and rates are compatible with the program’s level of complexity (Lapham et al., 2011).

Lastly, the program should focus on the competition strategy to be used for the initial award as well as for follow-on task orders and awards. This will help determine how to scope the contract or task order for each contract action. In some cases, the program would benefit from bundling a set of releases into a single contract action to minimize the number of contract activities during the development process. However, the program should balance this against the need to maintain continuous competition throughout the program lifecycle to keep rates low and receive the best value for products and services.

**Using a Contract Vehicle to Support Agile Program Structure**

As stated above, a services contract may represent a good strategy for a program seeking to acquire the skills and expertise of a developer to participate in a government-led Agile team. The program can pursue a separate stand-alone contract for Agile support services, or can consider leveraging an existing contract vehicle such as a GSA Schedule to acquire Agile support services on a task order basis. This strategy works well for a program that will need consistent Agile support to develop a single product, but is not recommended when pursuing a product-based contract, because the program would have to define requirements too far in the development process to gain the benefits of an Agile process (Modigliani & Chang, 2014). As illustrated in Figure 8, such a program would require consistent support throughout the development of several release cycles.

![Figure 8. Single-Award Services Contract](image)

A multiple-award IDIQ contract can allow a program to use several development contractors. This strategy would enable the program to maintain continuous competition for
future task orders and/or execute parallel development. Under this strategy, the program would award IDIQ contracts to two or more qualified vendors to compete on individual task orders, as illustrated in Figure 9. The program office would have to work closely with the contracting office to streamline contract timelines to enable rapid execution of task orders. This could be achieved by using standardized business practices, templates, and streamlined selection criteria. Past performance on task orders would become a weighted selection criterion for future work, further motivating contractor performance.

However, this strategy can also complicate integration and require increased resources to award and manage multiple contracts and developments (Lapham et al., 2011). To mitigate the integration risks of using two or more vendors, the government must dedicate time and effort to developing a rigorous architecture, interfaces, standards, and systems engineering processes. Each vendor should have active representation on the systems engineering Integrated Product Team to ensure a common understanding and maturation of these systems engineering elements throughout development. To foster coordination across vendors, the program should require the use of a common tool suite in the Request for Proposals process, and should also identify an initial set of required metrics each vendor must collect and report. In accordance with the contract, within the first 90 days of contract award, the vendors must submit to the program office an agreed-upon updated set of metrics proposed for review and approval.

If the program has reached a more mature stage of development with clearly defined releases, it may be feasible to execute product-based task orders. If requirements are dynamic and the program is in the initial stages of executing Agile, it would make more sense to use a service task order and compete the task orders for a set of releases.

**Summary**

The focus on iterative development and frequent capability deployments makes Agile an attractive option for many DoD IT acquisition programs, especially time-sensitive and
mission-critical systems. However, Agile differs so profoundly from traditional development practices that the DoD must overcome significant challenges to foster greater Agile adoption. The DoD cannot expect individual programs to tailor current acquisition processes on their own, because the complexities of the DoDI 5000.02 processes do not lend themselves to obvious solutions, let alone solutions that accommodate processes so fundamentally different from current DoD practices. Following the guidance offered in this paper would better equip programs to tailor the DoDI 5000.02 for Agile execution. As they face the next challenge of defining requirements in a way that meets rigorous JCIDS standards, programs can use the IT Box model outlined in this paper to enable the speed and flexibility required for Agile requirements. Lastly, programs can utilize the contracting strategies presented in this paper to acquire development support and utilize flexible contract vehicles that support Agile practices.

This paper has offered potential solutions to these key challenges in order to aid programs in laying a foundation for successful Agile implementation. As Agile adoption continues to take root and expand across programs, the DoD would benefit from additional guidance and training to ensure consistent and pervasive success in Agile IT acquisition.

References


Finding the “RITE” Acquisition Environment for Navy C2 Software

Captain George Galdorisi, USN (Ret.)—is director of the Corporate Strategy Group at SPAWAR Systems Center Pacific where he helps guide the Center’s strategic engagement. Prior to joining SSC Pacific, he completed a 30-year career as a naval aviator. His operational assignments included commanding officer tours of HSL-43, the Navy’s first operational LAMPS Mk III squadron; HSL-41, the LAMPS Mk III Fleet Replacement Squadron; USS Cleveland (LPD-7); and Amphibious Squadron Seven. His last operational assignment was as chief of staff for Cruiser-Destroyer Group Three, where he deployed to the Western Pacific and Arabian Gulf in the USS Carl Vinson and USS Abraham Lincoln. [george.galdorisi@navy.mil]

Amanda George—is a strategic analyst for SPAWAR Systems Center Pacific (SSC Pacific). Working with SSC Pacific’s Corporate Strategy Group, Amanda supports SSC Pacific’s programs in command and control, widgets and apps, and agile software development. Amanda joined SSC Pacific in 2010. She recently received her master’s degree in Pacific international affairs from the University of California, San Diego with a dual concentration in international politics and economics. She graduated magna cum laude from the University of California, San Diego with a bachelor’s degree in political science with a focus on international relations. [amanda.george@navy.mil]

Michael Morris—has over 30 years of experience working with C4I systems, including most recently working as the PEO C4I project lead/ RITE project manager. In addition to leading SSC Pacific’s RITE Project, Morris has extensive experience with the DoD Application Storefront. As the lead systems engineer for the DoD Application Storefront, he is responsible for monitoring the design and development of product architecture and process, standards, specification, and testing documentation to ensure all aspects of the Storefront contract are accomplished. [michael.a.morris4@navy.mil]

Abstract

The U.S. Navy has increasingly emphasized dominating the information battlespace as a key aspect of its warfighting strategy. As the Navy Strategy for Achieving Information Dominance states,

> whether characterized as intelligence, surveillance, reconnaissance, networks, communications, space, cyber, meteorology, oceanography, or electronic warfare, the Navy is inextricably and irreversibly dependent on information. Information provides a source of power but can also be an incapacitating weakness if not protected. Mastering the information domain is critical to the Navy’s future success. (U.S. Navy, 2012, p. 3)

A key aspect of mastering the information domain is mastering the Navy command and control (C2) systems that Navy operators rely on to assemble, organize, interpret, and analyze information. In order to agilely respond to a variety of different situations, Navy C2 systems must be agile themselves. The acquisition of Navy C2 systems has historically promoted stove-piped, single-mission systems that cannot react to the needs of today’s warfighters.

Space and Naval Warfare Systems Center Pacific (SSC Pacific) has collaborated with PEO C4I to devise and test an alternate acquisition process to address the challenges inherent in Navy C2 software acquisition. The new Rapid Integration and Test Environment (RITE) is a new lifecycle model of Navy C2 software that places increased emphasis on early and frequent software testing, as well as necessary software engineering practices at the source code level. RITE is a combination of commercial best-practice software engineering standards and processes; community process governance; contractual guidance; automated testing and report generation tools; and hosted development, test, and automated distribution facilities using a government and industry team agile acquisition model. Through the use of
RITE, SSC Pacific and PEO C4I have beta-tested a process that can enable streamlined acquisition of Navy C2 software.

Introduction

The United States has reached a possible strategic inflection point in which the recent international paradigm of the United States as a unipolar world power is shifting to one in which a number of near-peer competitors and non-state actors are increasingly gaining influence. Ronald O’Rourke, a specialist in naval affairs for the Congressional Research Service, has stated,

World events since late 2013 have led some observers to conclude that the international security environment is undergoing a shift from the familiar post-Cold War era of the last 20–25 years, also sometimes known as the unipolar moment (with the United States as the unipolar power), to a new and different strategic situation that features, among other things, renewed great power competition and challenges to elements of the U.S.-led international order that has operated since World War II. (O’Rourke, 2014, p. i)

Since 2013, the United States has seen a shift in the international strategic environment that has been focused primarily on three different factors: Russia’s aggression in Eastern Europe, China’s growing military modernization and actions in the East and South China Seas, and the increasing need to address non-state actions, including transnational crime and terrorist groups.

Russia is aggressively seeking a larger role in East Europe, a role that includes pushing for regional integration, as well as intervening militarily when it feels its domestic interests are threatened. As the director of National Intelligence, James Clapper (2015) stated, “Moscow is pushing for greater regional integration, pressing neighboring states to follow the example of Belarus and Kazakhstan and join the Moscow-led Eurasian Economic Union” (p. 17). Additionally, “in Ukraine, Russia has demonstrated its willingness to covertly use military and paramilitary forces in a neighboring state—a development that raises anxieties in states along Russia’s periphery” (Clapper, 2015, p. 18). In fact, “some observers trace the beginnings of the argued shift in strategic situations back to 2008” (O’Rourke, 2014, p. i) when “Russia invaded and occupied part of the former Soviet republic Georgia without provoking a strong cost-imposing response from the United States and its allies” (O’Rourke, 2014, p. 4). In relations with its Eastern European neighbors, Russia has been using “so-called ‘ambiguous warfare’ tactics” (O’Rourke, 2014, p. 9), where it has supplied military support—often including troops—but covertly. Taken in context of Russia’s recent actions, “the Ukrainian crisis has profoundly affected Russia’s relations with the West and will have far-reaching effects on Russia’s domestic politics, economic development, and foreign policy” (Clapper, 2015, p. 17).

The United States’ rebalance to the Asia-Pacific region has been initiated in large part due to China’s growing economic and military might, as well as its actions in the East and South China Seas. Clapper (2015) stated, “China will continue to pursue an active foreign policy—especially within the Asia Pacific—bolstered by increasing capabilities and its firm stance on East and South China Sea territorial disputes with rival claimants” (p. 19). In addition to maintaining an active foreign policy, like Russia, “China will probably seek to expand its economic role and outreach in the region, pursuing broader acceptance of its economic initiatives, including the Asia Infrastructure Investment Bank” (Clapper, 2015, p. 19). In pursuing its claims in the East and South China Sea, China has focused on a “so-called ‘salami-slicing’” (O’Rourke, 2014, p. 9) tactic in which it chips away at others’ claims to slowly extend its power.
As noted in the *National Security Strategy*, “The threat of catastrophic attacks against our homeland by terrorists has diminished but still persists” (Obama, 2015, p. 9). Moreover, potential terrorist adversaries “are not confined to a distinct country or region. Instead, they range from South Asia through the Middle East and into Africa” (Obama, 2015, p. 9). As Clapper (2015) stated, “Sunni violent extremists are gaining momentum and the number of Sunni violent extremist groups, members, and safe havens is greater than at any other point in history” (p. 4), so clearly, the threat is persistent. The DoD will continue to need to structure the military forces to address this threat in a variety of ways as terrorist attacks continue to evolve.

This shift in the international strategic environment has led the DoD to seek for new ways for the military to continue to support our national interests of security of the United States, its citizens, and U.S. allies and partners; a strong, innovative, and growing U.S. economy in an open international economic system that promotes opportunity and prosperity; respect for universal values at home and around the world; and a rules-based international order advanced by U.S. leadership that promotes peace, security, and opportunity through stronger cooperation to meet global challenges. (Obama, 2015, p. 1)

The Quadrennial Defense Review and the National Security Strategy outline the underpinnings of the defense strategy as it stands today. Based on this strategy, the United States will focus on the three following defense missions:

- **Protect the homeland**, to deter and defeat attacks on the United States and to support civil authorities in mitigating the effects of potential attacks and natural disasters. (DoD, 2014, p. v)
- **Build security globally**, in order to preserve regional stability, deter adversaries, support allies and partners, and cooperate with others to address common security challenges. (DoD, 2014, p. v)
- **Project power and win decisively**, to defeat aggression, disrupt and destroy terrorist networks, and provide humanitarian assistance and disaster relief. (DoD, 2014, p. v)

While each of these missions is necessary to preserve the national interests set out in the National Security Strategy, achieving all of these, given the shifting international strategic environment, is a tall order. To address the evolving security environment, the DoD introduced *The Defense Innovation Initiative* on November 15, 2014 (Hagel, 2014c).

This Defense Innovation Initiative (DII) has its roots in the offset strategies developed by national security professionals in the 1950s and the 1970s to ensure America’s military’s superiority. The first of these was President Eisenhower’s New Look in the 1950s, which prioritized nuclear deterrence. This was followed in the 1970s by the offset strategy of the Long-Range Research and Development Planning Program, which shaped future investments in leap-ahead capabilities such as standoff precision strike, stealth aircraft, wide-area surveillance, and networked forces. Under Secretary of Defense for Acquisition, Technology, and Logistics, Frank Kendall, explained that these two previous offset strategies yielded an impressive set of military capabilities, comprising a “revolution that we unleashed on the world in the first Gulf War” (Roulo, 2014). He went on to state that while the United States has continued to rely on this set of capabilities in the decades since, adversaries have had time and space to respond by building similar capabilities, which has spurred focus on the development of a Third Offset Strategy (Roulo, 2014).
Defense leaders—including Secretary of Defense Chuck Hagel and Deputy Secretary of Defense Robert Work—explain that the Third Offset Strategy is being driven by the growing risk to America’s continued technological superiority. In particular, Secretary Hagel (2014a) has stressed the threat posed by technology proliferation, noting that “disruptive technologies and destructive weapons once solely possessed by only advanced nations have proliferated widely, and are being sought or acquired by unsophisticated militaries and terrorist groups.” He also identified the threat from near-peer competitors China and Russia, stating that while the United States has been conducting stability operations for the past decade, China and Russia have been heavily investing in military modernization programs in order “to blunt our military’s technological edge” (Hagel, 2014b). In particular, Secretary Hagel emphasized that “they are … developing anti-ship, anti-air, counter-space, cyber, electronic warfare, and special operations capabilities that appear designed to counter traditional U.S. military advantages—in particular, our ability to project power to any region across the globe by surging aircraft, ships, troops, and supplies” (Hagel, 2014b). Budgetary constraints facing the DoD have made this threat environment even more challenging, limiting the Department’s ability to respond through an increase in the size of our military or simply outspending adversaries.

In response to this challenge to the United States’ technological and military superiority, the Third Offset Strategy—as instantiated in the Defense Innovation Initiative—seeks to put “the competitive advantage firmly in the hands of American power projection over the coming decades” (Hagel, 2014c). It will do so through several interrelated areas: a technology effort through the Long Range Research and Development Plan, leadership development practices, a new approach towards wargaming, operational concepts, and a continued focus on more efficient and effective business practices. As Secretary Hagel emphasized in assessing the previous two offset strategies, “The critical innovation was to apply and combine these new systems and technologies with new strategic operational concepts, in ways that enable the American military to avoid matching an adversary “tank-for-tank or soldier-for-soldier” (Hagel, 2014b).

A key concept of the DII is the focus on efficient and effective business practices. This area builds on a number of efforts the DoD has stood up in the last several years focused on streamlining both the oversight structure and business practices of the entire department. One central effort of this focus on efficient and effective business practices has been Under Secretary Kendall’s series of Better Buying Power strategies designed to reform the defense acquisition system to enable more efficient and rapid fielding of DoD technologies. On September 19, 2014, Under Secretary Kendall released an interim draft of the third instantiation of the Better Buying Power, Better Buying Power 3.0. Kendall (2014) noted in the release,

> Better Buying Power (BBP) is based on the principle that continuous improvement is the best approach to improving the performance of the defense acquisition enterprise. The evolution from BBP 1.0 to BBP 2.0 was based on the premise that emphasis would shift as initiatives were put in place, experience was accumulated, data was collected and analyzed, and conditions changed. BBP 3.0 continues that approach with a shift in emphasis toward achieving dominant capabilities through innovation and technical excellence. (p. 2)

BBP 3.0 is a key part of the DoD’s DII, in part because of their common concern that the technologies needed by military services are not being provided in a timely and efficient manner. The interim draft of BBP 3.0 states,
Underpinning BBP 3.0 is the growing concern that the United States’ technological superiority over potential adversaries is being threatened today in a way that we have not seen for decades. Our military today depends on a suite of dominant capabilities that originated in the ’70s and ’80s, has been enhanced and upgraded since, but has not fundamentally changed. This suite includes precision munitions, wide area surveillance systems, networked forces, and stealth technology. (Kendall, 2014, p. 2)

The concerns noted in the interim draft released in September are exactly aligned to the strategic concerns that underpin the DoD’s effort to create a third offset strategy.

BBP 3.0’s focuses on changing the culture and mindset in the defense acquisition community as well as ensuring that there are agile processes that defense acquisition experts can use to acquire and deliver cutting-edge capabilities. BBP 3.0 states,

One of the dominant characteristics of defense acquisition is its scope and complexity. There are no simple solutions to all the myriad problems acquisition professionals have to solve. There is no short “rule set” that will tell us all we need to know. (Kendall, 2014, p. 3)

This is true in part because the defense acquisition system is used to acquire a large variety of technologies. The same acquisition processes are used to acquire technologies that range from Navy Aircraft Carriers to command and control software systems, and Army tanks to cyber security constructs. Given the fluidity of the current and future security environment, there is no reason to expect that the need to acquire this wide variety of technologies will change any time soon. In fact, BBP 3.0 states,

Potential adversaries are modernizing at a significant rate, and they are responding rapidly to our development programs and fielded systems. This is true of peer, near-peer and even less capable potential adversaries. Our technology development and system designs must accommodate this reality. We must plan for likely responses to our designs, and we must be watchful and responsive ourselves to emerging threats. (Kendall, 2014, p. 4)

Thus, the DoD acquisition system must be able to accommodate the ability to agilely acquire and field a variety of systems.

The U.S. Navy—as well as the other services—relies upon the defense acquisition process to field its crucial command and control systems. While these systems do have a variety of “hard” physical components, much of the value of the system is based upon the “soft” software components that act as the brains of the systems. The software pieces of C2 systems—like Naval Integrated Tactical Environmental System–Next Generation (NITES-NEXT), Maritime Tactical Command and Control (MTC2), and Distributed Common Ground Station–Navy (DCGS-N)—are being updated continually in order to address a variety of the challenges that command and control systems face today.

Working with PEO C4I’s PMW-150, SSC Pacific has instituted the Rapid Integration and Testing Environment (RITE) to improve software development, testing and fielding. As Garcia explained, RITE is one of the strategic objectives that has been implemented “to support the Maritime C2 Roadmap” (2010, p. 17). He went on to note that “RITE is changing PMW 150’s software development methodology and modernizing the development process” (Garcia, 2010, p. 17). Utilizing the RITE will be a key component in enabling C2 software capabilities to be deployed more efficiently and effectively to the warfighter.
What is RITE?

RITE is a combination of commercial best-practice software engineering standards and processes, community process governance, contractual guidance, automated testing and report generation tools, and hosted development, test and automated distribution facilities using a government and industry team agile acquisition model. RITE is composed of four pillars—contracts, infrastructure, processes, and organization—which work together to facilitate the development and distribution of Navy C2 systems.

RITE was first initiated as a response to PEO challenges in managing, producing, and fielding C2 systems. The initial effort focused on finding ways was to improve product quality, move to competitive contract strategy, streamline acquisition cycle to meet aggressive timelines, and maximize release confidence and integration flexibility. The idea was to recognize that there is no silver bullet to solve these problems, but the approach should instead start with the most basic item: source code management. The key behind right is focus on improving component quality before delivery to decrease the average time needed to correct defects. Figure 1 illustrates how the average time needed to correct a defect increases based on how far it is from the initial code stage.

![Figure 1. Average Time to Correct Defect](image)

As Figure 1 illustrates, the closer the defect is identified to the code stage, the less time is needed to fix the problem. Together, the four RITE pillars work together to create an environment that makes it possible to manage software development as close to the source as possible.

**RITE Contracts**

The RITE contract pillar focuses on providing specific language that can be used in any contract that encourages a collaborative relationship. The RITE, as implemented by SSC Pacific, does not require a special contract vehicle type and, to date, has been delivered using existing Cost Plus Fixed-Fee (CPFF) contracts. Overall, CPFF has proven to provide the most cost-effective vehicle to support the research and developmental nature of the work that SSC Pacific performs. SSC Pacific has found that system engineering, software development, and test and integration support requirements are best met when the vendor is able to support evolving program requirements and to take advantage of changes in technology. SSC Pacific has inserted specific contract language within individual contracts/task orders and has tailored selected Data Item Descriptions (DIDs) to ensure that
the vendor is contractually obligated to work in close collaboration with SSC Pacific to implement the RITE model.

The objectives of the RITE contract pillars are focused on the five aspects: ask, use, receive, verify, and ensure. First, SSC Pacific uses contract language to ask for what the project needs, including specifics and requirements. Second, RITE is focused on using developers based on performance measures. The idea is to reduce the reliance on a single vendor, reduce barriers to entry, and decrease switching costs. RITE uses performance measures to ensure that there are increased choices at a competitive price throughout the development system. Third, RITE contracts ensure that SSC Pacific receives all of what it pays for. To that end, RITE requires that SSC Pacific receives buildable source code only as delivers—this includes source code for science and technology projects. Fourth, with RITE, SSC Pacific verifies that it receives what was promised through a rigorous software quality assurance. Finally, RITE uses enterprise repository contents and buildable source code operations to ensure that SSC Pacific can reproduce what it paid for. Integrating all five of these aspects into a single contract requires a great deal of planning up front, but pays big dividends over the life of the project.

The key feature in making RITE contracts as effective as they can be is the government—contractor relationship that is formed. SSC Pacific engages early to help determine the correct contract type. This engagement includes providing assessed risk of integration development or adapter software development and mitigating risk using RITE’s program comprehensive and coordinated Software Developer’s Kit (SDK). It is also key that all parties provide necessary documentation to ensure full and open competition. SSC Strive to create documentation consistent with the performance-based contracting approach. SSC Pacific will also advocate for the government’s unlimited rights in the technical data and software delivered by the contractor. Finally, RITE allows the government to use Multiple-Award Contracts—creating a truly competitive award environment.

**RITE Processes**

RITE processes include source code analysis tools to provide better cost estimates at source code level, internal inspection of code to reveal the state of the system from a software engineering level and software complexity, dependencies and coverage, and incorporation of automated test tools, able to reduce the time required to run a large number of test cases and increase the number of test events completed in less time.

RITE processes rely upon an iterative “sprint” model of development and delivery. Figure 2 illustrates the RITE development model.
As Figure 2 illustrates, each requirement is first broken down into a number of tasks. These tasks are then accomplished in a series of “sprints” with each sprint including a full iterative cycle of the product. As the iteration detail shows, each of the sprints includes the full cycle of examination of detailed requirements, design and analysis, implementation and testing, quality assurance and acceptance testing, and software deployment, and finishes up with an evaluation of the deployed software and a prioritization of tasks for the next sprint.

As Figure 3 shows, the sprint model allows for the product to be released in short bursts to the customer, rather than waiting for release only at the final Milestone C.
The release of C2 software products as the development and demonstration phase progresses will get the C2 capability into the warfighter’s operations quickly, as well as enabling much shorter testing time, as the problems are found during the product build cycle, rather than after it.

RITE Infrastructure

RITE infrastructure relies upon a centralized repository to enhance project communication and collaboration. The centralized repository creates a framework for software distribution (i.e., an application store), documentation library, development areas, software testing tools and data, and centralized software configuration management. Essentially, RITE creates a Distributed Development Environment (DDE), which is a virtual collaborative environment that spans multiple organizations and/or multiple physical locations. As Garcia (2010) has described, “in a DDE, project members share ideas, information and resources, and actively collaborate to achieve a common goal. The primary advantage of DDE is availability of resources and access to software development tools from different locations” (p. 17). Figure 4 details the RITE central repository infrastructure.
Figure 4. RITE Central Repository Infrastructure

As Figure 4 illustrates, the C4ISR developers, each of the testing labs, and the operational users are all connected to a central repository where the C2 applications, software artifacts, and automation tolls are housed.

Having a single infrastructure with a central repository greatly aids in the effort to maintain configuration management (CM). RITE includes management of configuration management and planning. This allows the parties to produce and adjust the CM plan as needed as well as ensuring timeliness of the lifecycle. This centralized repository also helps in identifying artifacts to be under CM control for functional, allocated, and product baselines. Enabling configuration control allows for baseline, change, and release management. In addition, various control boards are established to oversee overall CM. The central repository also allows for a designated project lead to maintain complete access control. The central repository will also store reports on artifacts as to their status in lifecycle, release state, configuration, interfaces, and physical environment as well as reports on who has accessed the different pieces, when the pieces were accessed and the delivery schedule. The final piece of CM in the central repository includes validation that all artifacts are held, available, and in proper status as well as the functional configuration audit (FCA) which documents whether the system does what it is proclaimed to do.

RITE Organization

In order to ensure that RITE is implemented effectively, it must be organized with deliberation. The organization managing the product must be a carefully selected team of people who all understand their jobs up front, and is organized for success. Figure 5 details the ideal organization for a RITE project.
Figure 5. Ideal Organization for RITE Project

As Figure 5 illustrates, the government ownership of the project and process is a key aspect of ensuring that the RITE project is successful. In particular, it is important for the project manager, product owner, and at least part of the development team to be government employees. The utilization of contractors is a key aspect of RITE, as it is unrealistic to expect that the government will be able to supply all the qualified personnel needed for the software engineering required by RITE. In addition, contractors can often offer other key skills that may not be present in the government workforce. Utilizing a mix of government employees and contractors allows for the project to reap the benefits of each group’s unique skill set.

RITE in Action

SSC Pacific and PMW-150 have already used RITE with a variety of programs. The past record of RITE demonstrates that it can provide significant cost and time savings. The agile nature of the software development institutionalizes end user input all along the development process, and also gets the capability into the end user’s hands much earlier. Using this agile process enables quick changes to current C2 programs of record, which has the potential to give the Navy an asymmetric C2 advantage over adversaries with outdated systems. Finally, the RITE process can—and should—be utilized by the entire DoD as all services grapple with the challenge of fielding twenty-first-century technology with a twentieth-century acquisition system.

References


Use of Automated Testing to Facilitate Affordable Design of Military Systems

Valdis Berzins—is a professor of computer science at the Naval Postgraduate School. His research interests include software engineering, software architecture, reliability, computer-aided design, and software evolution. His work includes software testing, reuse, automatic software generation, architecture, requirements, prototyping, re-engineering, specification languages, and engineering databases. Dr. Berzins received BS, MS, EE, and PhD degrees from MIT and has been on the faculty at the University of Texas and the University of Minnesota. He has developed several specification languages, software tools for computer-aided software design, and fundamental theory of software merging. [berzins@nps.edu]

Brian Womble—is the deputy for open architecture in the Office of the Deputy Assistant Secretary of the Navy for Research, Development, Test, and Evaluation. He spent the first half of his career working as a system developer in the telecommunications industry in Dallas, TX. In 2001, Womble joined Lockheed Martin in Manassas, VA, on Open Architecture efforts within the U.S. Navy Submarine ARCI program. He joined the Navy civilian workforce in 2009 and is now leading efforts to transition the Naval Enterprise to Open Systems Architecture through an Open Business Model. [brian.womble@navy.mil]

Abstract

Efforts to develop and implement automated test capability within the Department of Defense have resulted in the development of a number of tools. Literature from 2007 references use of automated testing to reduce the design cycle time of software, and it has been noted as one of several key components included as part of the more comprehensive plan for transforming the business and technical approaches to become more responsive to Fleet readiness requirements with the goal of providing more agile, integrated capabilities for the Navy by increasing supportability, standardization, system interoperability, network security, and Joint alignment.

This paper describes efforts to implement software systems automated test capability and an analysis of the results of the effort. The paper examines how well the automated test capability performed and evaluates the impact on system development time compared to systems developed using more traditional methods. In addition, a review of lessons learned and recommendations for further enhancements are discussed.

Overview: The Testing Challenge

Infinity Is a Big Place

The common slogan is that testing can demonstrate the presence of errors, but cannot guarantee their absence. This is valid in almost all situations. The exception is software services, for which the input space is small enough for exhaustive testing. Inputs include input parameters, initial states, and data read from files or input data streams. Practical limits for exhaustive testing are roughly one 32 bit input, which would require about 4 billion test cases. If we assume a 4GHz processor that takes 10 clock cycles per instruction on the average, that would take 10 seconds times the number of instructions

\[ \frac{4 \times 10^9}{4 \times 10^9 \text{ instructions}} \times \frac{10 \text{ seconds}}{10 \times 10^9 \text{ instructions}} = \frac{1}{10^9} \text{ seconds} \]

\[ \approx 1 \text{ minute} \]

Therefore, exhaustive testing is infeasible for all but the smallest systems. However, automated testing can be effective in finding many errors, even if not all possible inputs are tested.

1 The views presented is this paper are those of the authors and do not necessarily represent the views of DoD or its components.
needed to execute the service. For a service with a 32 bit input, that takes 1,000 instructions per execution—time for exhaustive testing would be about three hours; if the service needs a million instructions per execution, time for exhaustive testing becomes about four months.

For most practical services, the size of the input space far exceeds those estimated above, and in many cases, it may be unbounded, or bounded only by the capacity of the hardware. Since the size of the input space is exponential in the length of the input, this gets ridiculously intractable very fast: For services with just two 32-bit inputs, multiply the exhaustive testing time estimates above by 4 billion. Many practical services have much larger input spaces than that. Even though real computing hardware consists of finite state machines, the number of possible states in our machines is so large that it is not practically distinguishable from infinity.

There Will Be Bugs Left Behind

A consequence of the above analysis is that essentially all practical software systems are delivered with remaining imperfections.

However, not all faults are created equal: Every fault results in a failure for some subset of the input space. Since these spaces are finite, although very large, the failure spaces can be measured by the number of points in the failure space, or by the fraction of the input space occupied by the failure space. The latter fraction can be interpreted as the failure probability or failure rate associated with a given fault. Although the exact numbers involved are generally too large to be determined exactly, they can be estimated within given error tolerances by statistical sampling methods, and they can be used as a conceptual tool for classifying faults according to the associated failure rates.

The faults with the highest failure rate are those that produce a failure for all possible inputs. These are the faults with the highest impact on quality of service, and fortunately they are also the easiest to detect, since any single test case will detect all of them simultaneously. Faults with lower failure rates are increasingly difficult to detect via black box testing.

At the other extreme are the single-point failures: faults that result in a failure for only a single point in the input space, and produce correct results for all other input values. This category of failures is statistically invisible in practice, since a number of test cases close to exhaustive testing would be required to have an appreciable probability of detecting them in the absence of additional information about the fault.

Critical Bugs Are Must Fix

Failures are also not created equal: Some failures have more severe consequences than others. The critical bugs are those with the most severe consequences. Exposure is a combination of failure rate and severity of consequences. Severity of consequences can be measured with an abstraction known as “risk,” which can be estimated subjectively as a function of severity and failure rate, which can be interpreted as the expected likelihood of failure and relative overall cost to the Enterprise. In DoD contexts, severity of consequences has additional dimensions that include human injury and loss of life, in addition to financial loss. Risk is widely characterized as

\[ R(f) (s, fr). \]  \hspace{1cm} (1)

MIL-STD-882E (DoD, 2012) recognizes that severities and failure rates for particular hazards are rarely known exactly, and provides an approximate method for ranking hazards by degree of risk exposure that depend on subjective qualitative assessment of severity and failure rates based on informally defined ranges. The standard also provides guidance on
the level of authority needed to accept residual risks in each category of risk exposure. The practical result of this guidance is that the highest risks must be mitigated by measures that reduce the severity, failure rate, or both. This includes fixing the known faults with the highest risk exposure.

The weak point of current practice involves the word “known” in the previous sentence—all too often, faults become known only after at least one associated failure has occurred, along with associated undesired consequences. We would prefer faults with potentially severe consequences to be detected prior to fielding and actual occurrence of any failures due to those faults. This can be done by risk-based testing, which is done by automated testing whose intensity is determined by a risk analysis (Berzins, 2014).

**Finding Critical Bugs Requires Cheating**

Critical bugs may be statistically invisible. This is often the case for faults that are deliberately placed in the code by malicious insiders, such as Easter eggs and back doors. Such malicious additions to code are likely to be explicitly designed to produce statistically invisible single-point failures (see There Will Be Bugs Left Behind section). Such items are likely to be placed in services whose input spaces are much larger than the maximum size feasible for exhaustive testing.

For example, an Easter egg could be placed in a spreadsheet that would only be activated if a particular key was entered in a particular cell and all of the other cells were empty. Suppose that a cell can hold 10 characters and the spreadsheet can have 100 rows and 100 columns (most spreadsheets can accept much more data than this). If the testers know that only one cell is non-empty, then the number of possible input states is $10,000 \times 2^{80} > 10^{28}$. The probability of detecting the Easter egg by black box testing using 4 billion random test cases (roughly the largest practical amount) is approximately $4 \times 10^{-19}$, which is less than the likelihood of winning the grand prize in your favorite lottery twice in a row. Without knowing anything about the pattern that triggers the Easter egg, the number of possible input states becomes $2^{800,000}$ and the probability of detection by intensive black box testing would be less than $10^{-239,990}$, which is less than the chances of winning the grand prize in the lottery every day for the next $10^{26,640}$ years, even if we increase the number of test cases to the number that could be executed by all the computers in the world working for a century. To put that in perspective, the length of the winning streak is about $10^{26,640}$ times the age of our universe. This example is intended to illustrate that “statistically invisible” means “impossible to detect by black-box testing.”

Clear box testing can do better than black box testing in such cases, by using traditional coverage criteria, such as ensuring that every statement in the program has been executed for at least one test case. Running the usual test cases and keeping track of which statements have been executed, which can be done via instrumentation capabilities optionally provided by many compilers, will expose the rare paths in the code. Difficulties associated with covering the remaining statements include finding test inputs that exercise particular statements and determining whether remaining statements are in fact unreachable code. Although both problems are algorithmically unsolvable in the general case, for the kind of code encountered in practice, constraint solvers can succeed in synthesizing suitable test inputs for the majority of the cases and for identifying some of the unreachable code. The remaining code can be small enough to be singled out for human inspection.

**Software and Hardware Are Never Finished**

Successful systems always have long lists of pending change requests, including repairs for discovered faults and requested enhancements to functionality. In the Navy, such
changes are typically implemented in technology upgrade cycles that occur every two or every four years, depending on the program. Each change has the possibility of introducing new bugs into the system, and each new release must therefore be re-tested. This implies that tests must be repeated many times during the lifetime of a typical system.

Improving Affordability by Automated Testing

In practice, testing accounts for a substantial fraction of the cost of developing each new release. An online game called the Massive Multiplayer Online War-game Leveraging the Internet (MMOWGLI) was run in two rounds during 2013. This web-based game involves large numbers of distributed players who interacted to encourage innovative thinking via crowd-sourcing, generate ideas for solving problems, and plan actions that carry out identified solutions.

The second round of the MMOWGLI game addressed the issue of reducing cost without reducing system quality or capability. Several of the highest-ranked action plans produced by the game included automated testing and retesting as part of the strategy for affordably ensuring system quality (Schmidt, 2014).

Testing Is a Design Requirement

Hardware Is Designed With Test Points

Computing hardware, particularly integrated circuits, is designed to include special interfaces for testing. The purpose of these interfaces is to provide observability and controllability of internal states of the circuit. These are necessary because internal points on the chip are physically inaccessible and because the yield of manufacturing processes is less than 100%. Uncontrollable variations in manufacturing conditions, such as imperfect alignment of lithography masks and imperfect printing and etching due to dust particles in the air and working fluids, result in fabricated geometries that deviate from the ideal design. Some fraction of these result in chips that do not behave as designed. Successful sales depend on rapid acceptance testing that separates the functional chips that can be sold from the damaged ones that must be discarded.

Testing of digital hardware is easier than testing software for many reasons, including the following:

- *Uniform state representation*. For the purposes of testing, circuit state can be usefully represented as fixed-length bit vectors. When the circuit is in testing mode, all internal state cells are configured into a long shift register that can be sequentially output through the pins for observability and input from the pins for controllability. This enables open loop testing, where each test sets the internal state to a specified value, executes chosen operations, and then the internal state is read out for analysis. This avoids the problem of finding input sequences that will drive a possibly faulty circuit into prescribed initial test states, reduces time to design test cases, and speeds up the actual execution of the tests. In contrast, software states are typically sensitive to the meaning of the data, which varies widely between applications. This precludes a one-size-fits-all solution to observability and controllability of internal states.

- *Known expected outputs*. Since hardware tests are looking for deviations from the designed behavior, expected outputs can be derived using a uniform and conceptually simple process: Simulate the logical design on the test inputs, and calculate the expected results. This process is typically
completely automated. Since software tests are looking for design faults, finding expected outputs is a much harder problem that does not have an easy, uniform solution and generally requires human creativity for each new application.

- **Effective error models.** The processes that introduce manufacturing defects are well understood and produce defects that are easy to characterize. The most common defects are voids in conductors (manufactured circuits lack connections that are present in the design) and bridging between adjacent conductors (manufactured circuits have extra connections that are not present in the design). The vast majority of hardware faults can be effectively detected by test sets that expose all single stuck-at faults, and practical algorithms for automatically constructing such test sets are known.

Software has much more complex failure patterns, and complete test sets for detecting such patterns are not algorithmically computable in the general case.\(^2\) Observability and controllability for internal states of software are discussed in the next section.

**Architecture Assessment for Testability**

Software architectures can have a great impact on the effort required for system testing and the effort required to employ automated testing. Recent efforts by the Technical Reference Frameworks Working Group sponsored by ASN RDA have developed a structured set of testability levels to help assess these effects, shown in Table 1.

\(^2\) This is a consequence of Rice’s theorem, a well-known undecidability property.
Table 1. Testability Levels

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Inadequate</td>
</tr>
<tr>
<td>1</td>
<td>Syntactic</td>
</tr>
<tr>
<td>2</td>
<td>Semantic</td>
</tr>
<tr>
<td>3</td>
<td>Robust</td>
</tr>
<tr>
<td>4</td>
<td>Observable</td>
</tr>
<tr>
<td>5</td>
<td>Measurable</td>
</tr>
<tr>
<td>6</td>
<td>Decidable</td>
</tr>
<tr>
<td>7</td>
<td>Unbounded</td>
</tr>
</tbody>
</table>

Each level incorporates the requirements for all lower levels. The rationale for Table 1 can be explained as follows:

1. The externally observable behavior of a system consists of the services it provides to other systems. To enable independent testing of the system and its components, at a minimum the names of those services, the types of input data that each requires, and the types of data that each produces must be available to the testing team so that the services can be invoked by associated testing procedures. At level 1, this information should be specified as part of the system architecture to enable testing at each granularity level.

2. Although validation testing can be done at level 1, by relying on stakeholder review of each test output to judge adequacy of demonstrated behavior, verification testing requires level 2. Level 2 requires the architecture to include documented requirements for each service that are sufficiently precise to enable the testing team to make pass/fail decisions regarding the test outputs for each test case and service required at this level. This includes precise definitions of the properties of the real world that affect the requirements but may not be directly observable by the software. For example, a safety requirement in an aircraft control system typically specifies a minimum acceptable separation between aircraft. The requirement applies to the actual physical separation between the planes, rather than to the data visible to the control system software, which may differ because they are derived from sensors that can fail or produce inaccurate results. Level 2 may require human judgment for pass/fail decisions on test outputs, but those judgments can be made by the testing team, without requiring stakeholder participation in every test.
3. Level 3 includes documentation of all constraints, restrictions, and exceptional conditions associated with each service under test, not just the expected normal case behavior. This information is needed to check robustness of system operation, and can be used as a guideline for designing test cases focused on this issue. Level 3 implies complete coverage of the requirements, including both what the system is required to do and what it is required to avoid doing, regardless of whether inputs are in a “reasonable” range. For example, architectures at testability level 3 should include requirements on system input that would guard against SQL injection attacks; this information would support development of test cases that check what system behavior would result from such attacks.

4. At level 4, all system attributes relevant to checking the requirements are observable via standardized interfaces that are part of the architecture. This enables the software to be tested without modification by the test team, for example, without the need to manually add instrumentation code, and it enables test cases and test scripts to be portable across development and testing environments. Instrumentation is an issue in modern designs that use information hiding and object-oriented structures to limit access to internal system states. Access to some of these attributes may be needed for testing purposes, although they may not be needed during system operation, and their presence during system operation may not always be desirable due to the possibility of introducing cyber vulnerabilities. Therefore the testing interfaces may be excluded from the fielded version of the system, but they must conform to the documented standards whenever they are present. If instrumentation code is needed, architectures with testability level 4 include clearly documented standards for those testing interfaces. Ideally those standards enable automated instrumentation of the code in a repeatable manner that does not require human coding effort.

5. At level 5, all requirements are defined precisely enough so that all pass/fail decisions can be made based on defined criteria and measurement methods, without any expert human judgment required. Level 5 may still require human effort to apply the criteria, but they must be repeatable by anyone following a detailed written procedure.

6. At level 6, all requirements are defined precisely enough so that all pass/fail decisions for test cases can be made automatically, by software, firmware, or hardware. This implies that execution of test cases and assessment of test results can be completely automated at this level. Such automated tests can be repeated quickly at minimal cost, although up front setup costs for creating the automated decision procedures would be required. Level 6 differs from level 5 in that all of the criteria and measurement methods have been defined down to the granularity of basic operations that can be automated. Auxiliary instrumentation and communication links for physical attributes may be needed to support automated tests. For example, information from onboard GPS receivers may be needed to test separation between planes for the case in 2 above.

7. At level 7, test inputs can be randomly generated at need, based on probability distributions called operational profiles, and all pass/fail decisions can be automated. The difference between levels 6 and 7 is that 6 can be met via a list of expected outputs for a fixed set of test cases, while 7 requires a general pass/fail checking procedure that works for all possible inputs. This
level implies that additional test cases can be created without any human effort, which makes the marginal cost of additional test cases very small. At this level, very large test samples are affordable. This implies very high degrees of statistical confidence in system reliability can be achieved relative to the system workloads characterized by the operational profile distributions.

An additional test capability that can be provided at level 6 or 7 is built-in-test (BIT), which means that completely automated test capabilities are integrated into the deployed system, and can be invoked out in the field. Such capabilities may be used to ensure that all systems are functional prior to a mission or to recover from some types of equipment failures. If included in the architecture, such capabilities should include procedures for remedial action and be targeted at the most frequent expected failures. For example, a well-designed built-in test should be capable of diagnosing which part has failed, and issue instructions regarding what needs replacing and how to do it, or automatically switch to a backup system and issue a warning about the degraded status of the system. A very simple example of such a built-in test capability is a warning that a battery needs replacement or recharging, based on internal sensors. The benefit of including such capabilities is the ability to recover from some system failures out in the field, without the cost and delay of returning equipment to a home base for repair.

It is not necessarily useful to require every system service to have automated or built-in-test capabilities. Automated testing is generally beneficial only if the tests will be repeated often enough so that reduced marginal costs outweigh the extra setup cost of developing the automated test capabilities for each requirement. This is illustrated further in the following section.

An additional situation where automated testing is beneficial is in mitigating severe system risks, where high statistical confidence in particular system properties is required. That generally requires large numbers of randomly chosen test cases, which becomes affordable only at testability level 7. For example, guidance in MIL-STD-882-D (DoD, 2000) suggests that failure rates for mishaps with catastrophic consequences should not exceed $10^{-6}$, which requires roughly 20 million test cases if the probability of sampling error leading to a false positive conclusion must also be no more than $10^{-6}$. See Berzins and Dailey (2009) for details on determining the number of test cases required to meet given statistical confidence levels.

Experience With Automated Testing

Space and Naval Warfare Systems Center Pacific’s Command and Intelligence Systems Division of the Command and Control Department is the author of a software development and acquisition initiative that is gaining momentum across the Navy and DoD. This initiative is no new big bang/silver bullet; it simply focuses on lowering the cost and risk of government-developed software by demanding closer government control of the baseline, focusing testing where needed, and streamlining processes to deliver capability faster by relying on agile development methods.

Historically, DoD software was developed utilizing Waterfall or Spiral development methods. A prime contractor would be awarded the contract to go and build applications, integrate them, and return once completed, for a major Development Test (DT). This process typically lasted anywhere from 12–36 months, and after testing, as much as 48 months, before a requested capability or version of software was fielded. Often the software would be released reliant on hardware that would not be backward supported, causing additional issues with the installation and fielding of the applications. The DoD strove to
change this process and leverage the agile method to enable delivery of those applications faster.

Agile is a total paradigm shift for programs or development teams. Agile and Waterfall differ in many ways. A literature search reveals sources documenting many of the differences. Table 2 outlines several of the differences that are most relevant to DoD programs.

Table 2. Agile/Waterfall Compare and Contrast

| Agile                                                                 | Waterfall                                                                   |
|-----------------------------------------------------------------------|                                                                            |
| Allows for fluid requirements shifts and changes                      | Does not accommodate changes in requirements easily                        |
| Typically requires smaller teams of dedicated developers focused on smaller applications | Typically is based on larger teams working more at the system level        |
| Requires that work be time boxed into Sprints with a working product demonstrable at the end of each time box | Does not have to provide any working components until delivery at the end of the development cycle |

There are many other differences between the two methods, but the last one in Table 2 is intriguing and one to explore.

In most DoD programs that require software, the applications are built and tested at the contractor facility while development occurs. The tests, called Contractor Tests (CTs), are typically the contractor’s best guess as to how the products are going to be incorporated into a system and used in a workflow to support the user’s needs. When contractors have the ability to interact with the end users to get a good feel for the end use workflow, testing is better and the products typically work more efficiently, but are still subject to major defects based on architectural changes or dependent application modifications that might hamper the applications from working correctly.

From this point, once built and tested in the developers’ facility, the software is sent to the Integration Facility. The Rapid Integration and Test Environment (RITE) method infers that the Integration is conducted by a government facility or trusted agent, such as a Federally Funded Research and Development Center (FFRDC). Once the software is delivered to Integration, Integration Tests (ITs) are conducted. This process, according to the RITE method, is a graduated set of tests that begin with scanning and analysis of the source code delivered to the government-managed repository. The RITE Process requires source code be analyzed to ensure that it is built with suitable quality, as defined by the program manager and Integration Team. Source code scanning, a waterfall technique adopted for use by agile, and “White Box” type testing of source code is a fundamental change to the way that DoD integration activities have conducted tests in the past. The ability for DoD entities to now look down to the source code line that is the root cause for the defect and point remediation to that specific line of code is very powerful and extremely helpful in managing precious taxpayer dollars in the ownership of software intense systems.

---

3 Quality measurements are based on industry standards by referencing items such as SQALE, ISO, or IEEE documents as determined by the program manager.
The Integration Team then takes the software, once scanned and compiled, and begins deployment into a system development sandbox. These sandboxes allow the components to be “bolted” together and subsequently tested as dependent components, providing more of the end system execution environment. Automation testing at this stage is conducted based on CT artifacts that have been collected in previous deliveries that go into generating the Automation Regression Library. The beauty of this is that as system capability and applications mature, so does the automation regression library. Many would argue that automation and the development of scripting automated testing procedures is a heavy investment for many programs to undergo. It is true that the investment is substantial initially, but after the program has the foundation scripts generated, the investment in time is minimal to keep them updated delivery to delivery.

In some programs, our goal for automation was set to cover 65% of our testing via Automated Regression Scripts. As we quickly learned, this took approximately six months of dedicated, full-time effort for two personnel to script up the 65% of foundation scripts that we would utilize for a software intensive program, and then a dedicated 20% of two people’s time each month updating the Regression Library. With a productivity factor of 100 hours per month for a test engineer for a given Sprint, 20 of those hours would be used to do nothing more than update the Regression Library. The payoff is that using these scripts allowed a program with a code base of over 6 million SLOC to run over 3,500 test cases in a period of less than 10 hours. This same set of test cases was run manually by six test engineers 10 hours per day for 20 consecutive days. See Table 3 for a breakdown.

Table 3. Automation vs. Manual Test Comparison

<table>
<thead>
<tr>
<th>Method of Test</th>
<th>SLOC</th>
<th>Test Cases</th>
<th>Hours to Execute</th>
<th>Scripting Investment</th>
<th>Scripting Update Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated</td>
<td>6M SLOC</td>
<td>3500</td>
<td>10 Hours</td>
<td>1200 Hours</td>
<td>40 Hours</td>
</tr>
<tr>
<td>Manual</td>
<td>6M SLOC</td>
<td>3500</td>
<td>600 Hours</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

By reviewing Table 3, one can see the investment in scripting hours over the period of six months to get adequate regression suite of test.

Additional analysis shows that this investment is quickly paid off in two months or Sprints of dedicated testing that would have otherwise had to be done manually, and that the Return On that Investment (ROI) is 560 hours times the number of months or Sprints required after that. This enables the test team to repurpose the 560 hours that would have gone to repeat the regression test, adequate time to focus on new functionality or fix defects from one delivery to another. The ability to focus assets on new functionality facilitates the ability of the testing to cover a greater breadth of system capability and help decrease the possibility of major defects being released once the system is delivered. The goal is to achieve as close to 100% test coverage as possible.

On any delivery, the software produced typically satisfies anywhere from 1 to 200 requirements. This can require a proportional number of test cases to plow through in the space of the 30 days allocated to a Sprint. The only foreseeable way to test this new developed capability sufficiently, while at the same time ensuring that no previous capability has been broken as a result of the newly delivered code, requires dedication of hours from the test team, affordable only through automation. GUI Drivers, such as AutoIT, and test orchestration platforms, such as Test Complete, enable linking of automation scripts to generate predictive test workflows that mimic real world task execution. Testing can be planned based on certain requirements or mission components linked together to operate in certain environments. Additionally, increasing occurrences of executed instances or
increasing frequency of web service calls can generate load to test performance of a system.

One of the fundamental aspects of the RITE process is “focused testing.” Focused testing is accomplished in three phases. The first phase uses automated code quality analysis tools to scan software for Software Quality defects. These defects are relayed back to the developer upon identification for quick remediation. The second phase uses test scripts provided by the developer that are the foundation of automated functional testing to test services and component functionality build to build. The final phase uses human resources to manually test new features and develop automated scripts which will be used in future iterations of the focused testing process on successive software builds. Figure 1 shows the RITE recommended Testing Flow of Software through the Continuous Integration process.

![Figure 1. Continuous Integration Process Flow (Anchor A Continues the Process)](image)

Figure 1 also shows the flow of test-required dependencies as a software component evolves through integration. Rigorous testing requires a certain level of graduation from the lowest level, source code, to the higher components, and through to the System level. This graduated testing helps catch defects sooner, and enables fixes prior to fielding where the corresponding costs are much higher.

See Figure 2 for a sample Defect Fix Cost Chart. Figure 2 shows that as the Product Lifecycle progresses, the costs of fixing defects rise. This is typically where DoD programs suffer. Programs that subscribe to the RITE process, working from contracts that generate certain Quality Requirements for software to be developed and received, would lower the rates of defects being produced. In addition, the focused testing element of RITE also helps detect early on any defects that could find their way into the delivered source code. Finally, the use of an evolving automated regression suite, run repeatedly through the development and integration process, helps in decreasing the time spent in the model conducting a formal Development Test at the completion of a version release.
Conclusions

Automated testing has an important role to play in achieving affordable systems that reliably carry out their missions. This paper discusses automated testing, which primarily checks conformance of software behavior relative to system requirements. Other factors, such as quality of the requirements and the software architecture, are also relevant to system quality, and quality assurance techniques targeting those factors should be combined with automated testing for best results.

As explained in the Architecture Assessment for Testability section of this paper, effective automated testing depends on requirements that are both valid (capture the real needs of the stakeholders) and sufficiently well-defined to enable computing, whether particular test outputs conform to requirements or not. This is a challenge that will stress current requirements analysis processes, which typically produce natural language statements such as English descriptions of user needs. While such representations are needed for communication with people, they are insufficient by themselves for supporting automatic generation of test cases and automatic grading of test results. Natural language statements need to be augmented with more explicit representations that can support calculation of resulting truth values, such as logical assertions or the Object Constraint Language (OCL) associated with UML. This will require extra effort in requirements analysis, not only for coding requirements into these forms, but also for ensuring that the results are valid, and for refining the content of the requirements to provide sufficient definition detail and precision to carry out that encoding reliably. That extra effort is part of the initial investment needed to enable cost reduction by automated testing.

Complementary quality assurance processes are needed to ensure the quality of the software architecture and the subsystem requirements and specifications associated with that architecture. This is essential for affordably achieving reliability of large systems. A post-mortem analysis of software faults from the Voyager/Galileo programs found that the majority of the software faults were due to requirements and specification errors and misunderstanding of interfaces to external systems, not coding errors (Lutz, 1993). One of
the essential quality attributes for the requirements and specifications associated with a software architecture is a degree of consistency sufficient to enable harmonious interoperability between the subsystems specified in the architecture, because its absence leads to expensive system integration problems.

References
Panel 16. Managing Obligations Rates, Expenditures, and Internal Controls in the DoD

<table>
<thead>
<tr>
<th>Thursday, May 14, 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chair:</strong> Tom Mullins, Deputy Assistant Secretary of the Army for Plans, Programs, &amp; Resources</td>
</tr>
</tbody>
</table>
| *OSD’s Obligation & Expenditure Rate Goals: An Examination of the Factors Contributing to the Interference*  
  Robert Tremaine, DAU  
  Donna J. Seligman, DAU |
| *Contracting Processes, Internal Controls, and Procurement Fraud: A Knowledge Assessment*  
  Juanita M. Rendon, NPS  
  Rene G. Rendon, NPS |
| *Department of Defense Management of Unobligated Funds for Acquisition Programs*  
  William Fast, NPS |

Thomas E. Mullins—became the Deputy Assistant Secretary for Plans, Programs and Resources on November 13, 2005. He leads the planning, programming and execution of the Army’s Research, Development, Test and Evaluation and Procurement Appropriations totaling over $40B annually. His oversight includes serving as the co-chair of the Equipping and Sustaining Program Evaluation Groups that direct the Army’s long-term Sustaining and Acquisition strategies. Mr. Mullins provides fiscal and resource counsel to the Assistant Secretary of the Army (Acquisition, Logistics and Technology). He provides oversight and direction to Reserve and National Guard acquisition efforts.

Prior to his appointment as the Deputy Assistant Secretary, Mr. Mullins served as the Deputy Director for Plans, Programs and Resources under the Deputy Assistant Secretary for Plans, Programs and Resources. Prior to this assignment he was the Director of Maneuver Systems for the Deputy for Acquisition and Systems Management and the Program Executive Officer Representative for Ground Combat Systems. Mr. Mullins has also served as a project manager and as a member of the staff of the Under Secretary of Defense for Acquisition, Technology, and Logistics, Strategic and Tactical Systems, Land Warfare.

Mr. Mullins has over 42 years of federal service including 20 years as an active Army officer with expertise in program management and business and financial management. His civilian education includes a Bachelor of Science in Physics and a Bachelor of Science in Mathematics from the Central State University in Edmond, Oklahoma, and a Master of Science in Management from the Naval Postgraduate School. He has been a member of the civilian Army Acquisition Corps since 1992 and has achieved a Level III certification in Program Management. Awards include the Legion of Merit, Office of the Secretary of Defense Award for Excellence, and numerous Exceptional Performance awards.
OSD's Obligation & Expenditure Rate Goals: An Examination of the Factors Contributing to the Interference

Robert Tremaine—is associate dean, Outreach and Mission Assistance, Defense Acquisition University West Region. Col. Tremaine is a retired Air Force colonel and has over 26 years of experience in air, missile, and space weapons systems acquisitions. He holds a BS from the U.S. Air Force Academy and an MS from the Air Force Institute of Technology. Col Tremaine is Level III certified in both program management and systems planning, research, development, and engineering. [robert.tremaine@dau.mil]

Donna Seligman—is a management information systems manager at the Defense Acquisition University West Region. She has extensive experience with developing complex business knowledge applications, performing system analyses, and conducting research. She has also supported legacy code and reengineering efforts of other major business system applications. Seligman holds a BS in Information Decision Systems from San Diego State University. [donna.seligman@dau.mil]

Abstract
Managing DoD acquisition programs is a complicated process. The turbulence created by funding instability makes it even more difficult. To help program offices maintain their overall funding execution pace, the Office of the Secretary of Defense (OSD) instituted Obligation and Expenditure rate goals over two decades ago. Acquisition program managers have found it difficult to meet established Obligation and Expenditure rate goals. For the purposes of this study (sponsored by Nancy Spruill, director of Acquisition Resources and Analysis, Office of the Under Secretary of Defense for Acquisition, Technology and Logistics) and based on Defense Acquisition University and OSD subject matter expertise, the authors looked closely at potential causal factors that could be interfering with the achievement of these goals. Two hundred and twenty-nine DoD personnel responded to a comprehensive survey. The respondents were comprised of program office personnel (program managers [PMs], deputy PMs, budget and financial managers [FMs], and contracting officers); program executive officers (PEOs) and their chief financial officers; and a variety of senior OSD staff including Headquarters FM senior staff and Senior Acquisition Executive staff. The respondents were asked if they found metrics helpful in better meeting OSD goals as well as the use of any process improvements.

Introduction
In the months preceding this research effort, Nancy Spruill, director, Acquisition Resources and Analysis, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics (OUSD [AT&L]) solicited support from the Defense Acquisition University (DAU) to help uncover the causal factors that could be interfering with the attainment of the Office of the Secretary of Defense’s (OSD’s) Obligation and Expenditure rate goals. To learn more about the intervening obstacles, DAU, along with assistance from the OSD, developed a comprehensive survey that queried experienced and high-level Department of Defense (DoD) personnel involved in a weapon program’s decision chain. What we learned from the subsequent analysis confirmed several previous suspicions. The data also indicated the prevalence of more underlining perception variances among many of the factors that could be undermining program execution itself.
Research Methodology

Two hundred and twenty-nine DoD personnel responded to this survey. The respondents were comprised of program office personnel (program managers, deputy program managers, budget and financial managers, and contracting officers), program executive officers and their chief financial officers (CFOs), and a variety of senior staff at the OSD including Headquarter Financial Management (FM) senior staff and Senior Acquisition Executive (SAE) staff (Table 1). Because several functional areas saw lower response rates, a more detailed analysis of the causal factors was restricted to an aggregate sample size given the confidence levels required to draw any inferences or conclusions.

Table 1. Respondent Demographics

<table>
<thead>
<tr>
<th>ACAT LEVELS</th>
<th>Respondent Groups</th>
<th>TOTALS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Program Office(^2)</td>
<td>PEO(^3)</td>
</tr>
<tr>
<td>I</td>
<td>91</td>
<td>142</td>
</tr>
<tr>
<td>II</td>
<td>28</td>
<td>63</td>
</tr>
<tr>
<td>III</td>
<td>23</td>
<td>24</td>
</tr>
</tbody>
</table>

\(^1\) Includes sampling from all Components and several DoD agencies
\(^2\) Program Managers, Deputy Program Managers, Budget and Financial Managers (BFM), Deputy BFM, and Contracting Officers
\(^3\) Program Executive Officers (PEOs), Deputy PEOs and their Chief Financial Officers
\(^4\) Headquarter Financial Managers and Senior Acquisition Executive Staff

Respondents ranked the impact of 64 factors under nine categories (Figure 1). The researchers then assessed the rankings using a top box (TB) three methodology (i.e., the percentage of 5, 6, and 7 responses on a Likert-like scale from 1–7). Since the frequency of occurrence for some factors could also be contributing to the interference, the researchers included an additional selection (e.g., daily, weekly, monthly, etc.) to isolate any potential ignition areas.
Discussion

Factor Distribution

Figure 2 shows the distribution of all 64 factors assessed. Three factors reported an impact rating of two standard deviations above the mean (denoted by $+2\sigma$); six factors reported an impact rating of one standard deviation above the mean (denoted by $+1\sigma$); and 22 factors rose above an average impact rating (denoted by $\bar{x}$). The remaining 33 factors fell below the aggregate $\bar{x}$.

Nineteen of the 22 factors measured for frequency of occurrence resulted in an impact rating above 39%. Sometimes, just one occurrence appeared to have a significant impact.
The Causal Factors Rank Ordered

Table 2 lists the relative ranking of all 64 factors in the context of top box descending order to provide a comprehensive view of all factors, although the remaining discussion in this paper addresses the factors above \( \bar{x} \). Unrealistic Spend Plans (F10), also one of the factors assessed, is generally valued as a written forecast of a program’s funding needs and establishes Obligation and Expenditure projections. However, spend plans are subjected to so many real world eventualities that updating them becomes problematic in sustaining its forecasting value.
### Table 2. Impact Factor Ratings in Aggregate Descending Order

<table>
<thead>
<tr>
<th>Factors Rated by Adverse Impact</th>
<th>TB</th>
<th>( \bar{x} )</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 Late release of full obligation/budget authority due to Continuing Resolution Authority</td>
<td>69%</td>
<td>5.29</td>
<td>2.41</td>
</tr>
<tr>
<td>F2 Contract negotiations delays</td>
<td>67%</td>
<td>5.06</td>
<td>2.50</td>
</tr>
<tr>
<td>F3 Contract award delays</td>
<td>67%</td>
<td>5.00</td>
<td>2.50</td>
</tr>
<tr>
<td>F4 Shortage of Contracting Officers</td>
<td>64%</td>
<td>4.79</td>
<td>2.50</td>
</tr>
<tr>
<td>F5 Congressional mark</td>
<td>61%</td>
<td>4.97</td>
<td>2.66</td>
</tr>
<tr>
<td>F6 Contractor proposal prep delays</td>
<td>60%</td>
<td>4.97</td>
<td>2.50</td>
</tr>
<tr>
<td>F7 OSD directed RMD adjustment</td>
<td>56%</td>
<td>4.50</td>
<td>2.63</td>
</tr>
<tr>
<td>F8 RFP prep delays</td>
<td>57%</td>
<td>4.03</td>
<td>2.40</td>
</tr>
<tr>
<td>F9 Source selection delays</td>
<td>55%</td>
<td>4.44</td>
<td>2.63</td>
</tr>
<tr>
<td>F10 Unrealistic/overly optimistic spend plans</td>
<td>52%</td>
<td>4.30</td>
<td>2.44</td>
</tr>
<tr>
<td>F11 Changes in user requirements</td>
<td>51%</td>
<td>4.16</td>
<td>2.43</td>
</tr>
<tr>
<td>F12 Changes to program acquisition strategy</td>
<td>51%</td>
<td>4.41</td>
<td>2.52</td>
</tr>
<tr>
<td>F13 Changes in other stakeholder requirements</td>
<td>50%</td>
<td>4.32</td>
<td>2.34</td>
</tr>
<tr>
<td>F14 Preparing DAE level review and decision</td>
<td>50%</td>
<td>4.15</td>
<td>2.18</td>
</tr>
<tr>
<td>F15 Lack of decision authority at expected levels</td>
<td>60%</td>
<td>4.22</td>
<td>2.62</td>
</tr>
<tr>
<td>F16 Implementation of new OSD/Service policy</td>
<td>49%</td>
<td>4.20</td>
<td>2.59</td>
</tr>
<tr>
<td>F17 Component directed POM adjustment</td>
<td>49%</td>
<td>4.26</td>
<td>2.51</td>
</tr>
<tr>
<td>F18 Awaiting reprogramming action</td>
<td>49%</td>
<td>4.23</td>
<td>2.44</td>
</tr>
<tr>
<td>F19 Changes in user priorities</td>
<td>47%</td>
<td>4.00</td>
<td>2.38</td>
</tr>
<tr>
<td>F20 Realistic spend plans but risks materialized</td>
<td>45%</td>
<td>4.00</td>
<td>2.21</td>
</tr>
<tr>
<td>F21 Program delays from additional development, testing or other prerequisite events</td>
<td>44%</td>
<td>4.00</td>
<td>2.30</td>
</tr>
<tr>
<td>F22 DCMA administrative actions</td>
<td>44%</td>
<td>3.92</td>
<td>2.61</td>
</tr>
<tr>
<td>F23 Unplanned Congressional add to PE request</td>
<td>43%</td>
<td>3.90</td>
<td>2.74</td>
</tr>
<tr>
<td>F24 Use of unfinanced contract action delays</td>
<td>42%</td>
<td>3.73</td>
<td>2.50</td>
</tr>
<tr>
<td>F25 Expenditure contingent on hardware delivery</td>
<td>41%</td>
<td>3.92</td>
<td>2.41</td>
</tr>
<tr>
<td>F26 Loss of funding through reprogramming action to higher profit requirements to PEO portfolio</td>
<td>41%</td>
<td>3.89</td>
<td>2.46</td>
</tr>
<tr>
<td>F27 Lack of experience levels in key acquisition functional areas</td>
<td>40%</td>
<td>3.90</td>
<td>2.30</td>
</tr>
<tr>
<td>F28 Awaiting DAE level review and decision</td>
<td>40%</td>
<td>3.50</td>
<td>2.42</td>
</tr>
<tr>
<td>F29 Shortage of Cost Estimators</td>
<td>40%</td>
<td>3.87</td>
<td>2.37</td>
</tr>
<tr>
<td>F30 Shortage of Business/finance personnel</td>
<td>39%</td>
<td>3.73</td>
<td>2.30</td>
</tr>
<tr>
<td>F31 Programmatic conflicts between post and prime contractors</td>
<td>39%</td>
<td>3.86</td>
<td>2.32</td>
</tr>
<tr>
<td>F32 Preparing SAE/CAE level review and decision</td>
<td>37%</td>
<td>3.74</td>
<td>2.03</td>
</tr>
<tr>
<td>F33 Delays in contractor payment due to late invoices</td>
<td>37%</td>
<td>3.97</td>
<td>2.35</td>
</tr>
<tr>
<td>F34 Unobligated prior year funding not adequately factored</td>
<td>36%</td>
<td>3.57</td>
<td>2.23</td>
</tr>
<tr>
<td>F35 Component Comptroller Withhold</td>
<td>35%</td>
<td>3.58</td>
<td>2.34</td>
</tr>
<tr>
<td>F36 DCMA administrative actions</td>
<td>35%</td>
<td>3.42</td>
<td>2.36</td>
</tr>
<tr>
<td>F37 Reduction of contractor efforts</td>
<td>35%</td>
<td>3.47</td>
<td>2.73</td>
</tr>
<tr>
<td>F38 OSD Comptroller Withhold</td>
<td>34%</td>
<td>3.43</td>
<td>2.37</td>
</tr>
<tr>
<td>F39 Shortage of Tech/Eng/Test personnel</td>
<td>34%</td>
<td>3.51</td>
<td>2.17</td>
</tr>
<tr>
<td>F40 Shortage of Auditors</td>
<td>33%</td>
<td>3.17</td>
<td>2.43</td>
</tr>
<tr>
<td>F41 Slower burn rate than expected due to unfavorable SPI</td>
<td>33%</td>
<td>3.25</td>
<td>2.14</td>
</tr>
<tr>
<td>F42 Awaiting SAE/CAE level review and decision</td>
<td>32%</td>
<td>3.33</td>
<td>2.30</td>
</tr>
<tr>
<td>F43 SAE/CAE/Component directed reprogramming</td>
<td>32%</td>
<td>3.27</td>
<td>2.30</td>
</tr>
<tr>
<td>F44 Rejection</td>
<td>32%</td>
<td>3.16</td>
<td>2.46</td>
</tr>
<tr>
<td>F45 Changes in systems specs</td>
<td>31%</td>
<td>3.30</td>
<td>2.03</td>
</tr>
<tr>
<td>F46 Tenure of PM and others in key positions</td>
<td>31%</td>
<td>3.11</td>
<td>2.18</td>
</tr>
<tr>
<td>F47 Holding award/incentive fees in commitment for future obligation</td>
<td>20%</td>
<td>3.23</td>
<td>2.36</td>
</tr>
<tr>
<td>F48 Inadequate training</td>
<td>29%</td>
<td>3.30</td>
<td>2.13</td>
</tr>
<tr>
<td>F49 Shortage of Managers</td>
<td>29%</td>
<td>3.10</td>
<td>2.17</td>
</tr>
<tr>
<td>F50 Insufficient planned GO funding</td>
<td>27%</td>
<td>3.07</td>
<td>2.27</td>
</tr>
<tr>
<td>F51 Shortage of Staff</td>
<td>26%</td>
<td>2.99</td>
<td>2.12</td>
</tr>
<tr>
<td>F52 Contractor rework</td>
<td>26%</td>
<td>3.00</td>
<td>2.14</td>
</tr>
<tr>
<td>F53 Delayed payments for scheduling, earning fees, progress payments/performance based payments</td>
<td>25%</td>
<td>3.98</td>
<td>2.76</td>
</tr>
<tr>
<td>F54 Effect of contract type on outlay rates</td>
<td>24%</td>
<td>2.96</td>
<td>2.12</td>
</tr>
<tr>
<td>F55 Materiel/Systems Command Comptroller Withhold</td>
<td>24%</td>
<td>2.71</td>
<td>2.17</td>
</tr>
<tr>
<td>F56 Awaiting PEO level review and decision</td>
<td>24%</td>
<td>2.90</td>
<td>2.01</td>
</tr>
<tr>
<td>F57 Termination Liability</td>
<td>22%</td>
<td>2.72</td>
<td>2.17</td>
</tr>
<tr>
<td>F58 Insufficient workplace tools/apps</td>
<td>22%</td>
<td>2.92</td>
<td>2.01</td>
</tr>
<tr>
<td>F59 PEO directed programming</td>
<td>21%</td>
<td>2.93</td>
<td>2.10</td>
</tr>
<tr>
<td>F60 Slower burn rate than expected due to favorable CPI</td>
<td>21%</td>
<td>2.77</td>
<td>1.95</td>
</tr>
<tr>
<td>F61 PEO Withhold</td>
<td>20%</td>
<td>2.39</td>
<td>1.90</td>
</tr>
<tr>
<td>F62 Preparing PEO level review and decision</td>
<td>20%</td>
<td>2.96</td>
<td>1.53</td>
</tr>
<tr>
<td>F63 Production line issues</td>
<td>19%</td>
<td>2.82</td>
<td>2.08</td>
</tr>
<tr>
<td>F64 Labor disputes</td>
<td>10%</td>
<td>1.85</td>
<td>1.64</td>
</tr>
</tbody>
</table>

\( +2\sigma = 67\% \)

\( +1\sigma = 53\% \)

\( -1\sigma = 25\% \)
Factors and Respondent Groups

Figure 3 accounts for the 31 factors above the mean and by respondent group seen in Table 1. The 31 factors were the only ones further evaluated in this study unless a factor shifted above $\bar{x}$ after any further delineation (e.g., ACAT levels, military components, position, etc.). Unexpectedly, the individual factors showed widespread perception disparities among the respondent groups for the factors that fell below $+2\sigma$. After analyzing the specific individual factors among all the respondent groups, seven of the 31 factors had an unusually large $\sigma$. As a result of these conspicuous gaps, we turned to the qualitative data. We also watched for any strong correlations (e.g., correlation coefficients ($r$) > 0.7) to better understand the reasons for the differences, as well as the influence of any intervening and/or moderating factor couplings. The remaining discussion addresses the 31 impact factors in descending order from highest to lowest.

![Figure 3. Impact Factor Ratings Above $\bar{x}$ in Aggregate Descending Order With Respondent Group Low and High Ratings](image)

The Factors That Ranked Above $+2\sigma$

In Figure 3, late release of full obligation/budget authority due to Continuing Resolution Authority ($F^1$), Contract Negotiations Delays ($F^2$), and Contract Award Delays
all rose above $2\sigma$, where 67% or more of the respondents claimed they had the highest adverse impact of all factors measured. The occurrence of CRA had the most significant negative impact to Obligation and Expenditure rates. It also had one of the smallest variances ($\sigma$) among the respondent groups. Even with the expectation that CRA might prevail and the subsequent planning that followed for such a likely event, many PMs pointed to an overly conservative and slow internal vetting process posture that created additional obstacles in meeting OSD goals. Several PMs recommended the use of some sort of “CRA variable” to temporarily offset the consequences of CRA if the required funds were not released as originally projected. Next in rank order were contract negotiations and contract award delays. The respondents emphasized that the DoD could fix the problem more readily since unlike CRA, these factors were under internal control. When asked what could be done to reduce the adverse effects of all three factors, the respondents recommended the “inclusion of more risk mitigation into contract award planning, more realistic timelines, more realistic plans, greater funding stability, reduction in bureaucratic obstacles, more synchronized internal processes, and better aligned accounting systems.”

**The Factors That Ranked Above $+1\sigma$**

This next line of demarcation (Figure 3, factors $F^4$–$F^9$) included many contracting-related factors (i.e., Shortage of Contracting Officers [$F^4$], Contractor Proposal Prep Delays [$F^6$], RFP Prep Delays [$F^8$], and Source Selection [$F^9$]). Nearly all the factors showed the emergence of a more alarming $\sigma$ between the individual respondent groups—as high as 18% in one case (i.e., Proposal Prep Delays [$F^6$]). For this particular factor, PCOs reported the highest impact while PMs ranked it as the lowest. Senior staff cited that Shortage of Contracting Officers ($F^4$) created the highest impact, while PCOs reported it had the lowest impact. With a 7% $\sigma$, it was the lowest among all six factors in this grouping.

Given that six of the top nine factors were contract-specific factors that ranked above $+1\sigma$ (see Figure 3), it came as little surprise to see so many reinforcing comments surface:

- “Lack of experienced and qualified contract specialists”
- “Alarmingly low personnel qualified … many unsure/lack guidance and experience”
- Significantly stressed with overtime to complete all contracting actions prior to close of fiscal year”
- “Inadequate training … inordinate number of interns with very low experience in all career fields”
- “Lack of sufficient legal personnel trained in acquisition”
- “Loss in brain trust and skill to develop complete, clear SOWs using proactive contract language”
- “SOW writing and the teaching of SOW-writing classes is greatly left to contractors or support contractors resulting in unclear language”

The highest frequency of occurrence was also associated with contracting-related factors (Figure 3). By far, the aggregate respondents rated Shortage of Contracting Officers ($F^4$) as the single highest factor among all 22 factors measured for frequency. Because the contracting activity timeline generally has lengthy durations, any disruption appears to have an unmistakable impact on contract award. Shortage of Contracting Officers ($F^4$) was seen as having the most significant impact. Several respondents said that “multiple contracting actions were having compounding consequences.”
The two remaining factors above $+\sigma$, Congressional Marks (F5) and OSD Directed RMD Adjustment (F1), had very low frequency of occurrences but still reported a very high impact, similar to CRA. When combining these with F4, all three appear to be a strong antecedent force (or moderating factor) to the already time-consuming chain of contracting actions.

**The Factors That Ranked Above $\bar{x}$**

This final grouping (Figure 3, factors F10–F31) accounted for the remaining 22 impact factors. Perception polarities persisted especially between two respondent groups—senior staff outside the program office and PMs inside program offices. For PMs in every case except one (i.e., Component Directed POM Adjustment [F17]), the impact factors ranked well below $\bar{x}$. In sharp contrast, senior staff in every case except one (i.e., Component Directed POM Adjustment [F17]) stated the majority of top 31 factors had the largest impact or close to it among all respondent groups.

Even though the remaining impact factors above $\bar{x}$ are still significant, the researchers shifted the focus to the presence of any strong correlations since factor couplings could be having a moderating effect and require a closer look.

**The Factors That Correlate**

Table 3 summarizes the strongest and weakest factor correlations for all respondents queried. Several strong correlations surfaced for factors above $\bar{x}$. User Requirements (F11) and User Priorities (F19) were very strongly correlated. In three specific instances, two factors above $\bar{x}$ were very strongly correlated with three factors that fell below $\bar{x}$: Key Acquisition Experience (F27) and Inadequate Training (F48); Key Acquisition Experience (F27) and Tenure of PM and other Key Positions (F46); and DCMA Administration Actions (F36) and DCAA Administration Actions (F22). Three contract-related factors (F4, F8, and F9) showed weaker correlations than unexpected. Because a factor had a weak correlation does not mean it had any less importance, but any course of action intended to mitigate the presence of any impact factor strongly correlated with another should be weighed more heavily in any recommended action. For example, the turnover of PMs could be part of the experience quotient.

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strongest Correlation Coefficients</strong></td>
<td></td>
</tr>
<tr>
<td>F4</td>
<td>Experience and Training and Tenure: Key Acquisition Experience Levels &amp; F40 Inadequate Training</td>
</tr>
<tr>
<td>F5</td>
<td>Key Acquisition Experience Levels &amp; F40 Inadequate Training</td>
</tr>
<tr>
<td>F6</td>
<td>Key Acquisition Experience Levels &amp; F40 Tenure of PM &amp; Other Key Positions</td>
</tr>
<tr>
<td>F10</td>
<td>Administrative Actions: F10 DCMA &amp; F10 DCAA</td>
</tr>
<tr>
<td>F11</td>
<td>Changes in Program Context: F11 User Requirements &amp; F11 User Priorities</td>
</tr>
<tr>
<td>F12</td>
<td>F12 User Priorities &amp; F12 Stakeholder requirements</td>
</tr>
<tr>
<td>F13</td>
<td>Contract-related Activities: F13 Contract Award Delays &amp; F13 Contract Negotiations Delays</td>
</tr>
<tr>
<td>F14</td>
<td>F14 Contract Award Delays &amp; F14 Contract Negotiations Delays</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weakest Correlation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Late release of full obligation/budget authority due to CRA</td>
</tr>
<tr>
<td>F2</td>
<td>Shortage of Contracting Officers</td>
</tr>
<tr>
<td>F3</td>
<td>Congressional mark/Recession</td>
</tr>
<tr>
<td>F4</td>
<td>OSD Directed RMD Adjustment</td>
</tr>
<tr>
<td>F5</td>
<td>RFF prep delays</td>
</tr>
<tr>
<td>F6</td>
<td>Unreasonably optimistic spend plans</td>
</tr>
<tr>
<td>F7</td>
<td>Changes to program acquisition strategy</td>
</tr>
<tr>
<td>F8</td>
<td>Lack of decision authority</td>
</tr>
<tr>
<td>F9</td>
<td>Implementation of new OSG/Service policy</td>
</tr>
<tr>
<td>F10</td>
<td>Component DirectedPOM Adjustment</td>
</tr>
<tr>
<td>F11</td>
<td>Awaiting reprogramming action</td>
</tr>
<tr>
<td>F12</td>
<td>Realistic spend plans but risks materialized</td>
</tr>
<tr>
<td>F13</td>
<td>Program delays from prerequirement events</td>
</tr>
<tr>
<td>F14</td>
<td>Unplanned Congressional adds to FBD request</td>
</tr>
<tr>
<td>F15</td>
<td>Expenditure contingent on hardware delivery</td>
</tr>
<tr>
<td>F16</td>
<td>Funding Loss: reprogramming action to higher priority requirements to PEG portfolio</td>
</tr>
<tr>
<td>F17</td>
<td>Shortage of Business/finance personnel</td>
</tr>
<tr>
<td>F18</td>
<td>Programmatic conflicts between government and prime contractor</td>
</tr>
</tbody>
</table>

* The higher the % the stronger the direction and strength of the linear relationship between the variables

Factors # 1–3 ≥ $+20$; Factors # 4–9 ≥ $+10$; Factors # 10–31 ≥ $\bar{x}$
Factor Plotting

The researchers generated a scatter plot diagram (Figure 4) that punctuated how the 31 factors fluctuated between impact and frequency of occurrence. In some cases, the impact of certain factors occurred with low frequencies of occurrence. In other cases, the frequency of occurrence compounded the impacts.

The research data was rebased to a Likert-like scale for plotting the frequency and adverse impact response averages. The researchers included Factors F29–F31 in Figure 4 because they only fall slightly below $\bar{x}$.

![Figure 4. Scatter Plot of Impact Factors With Frequency](image)

For the relationships that were co-linear (e.g., the most strongly correlated depicted in Table 3), the researchers explored whether they also behaved as strong predictors across the sample population. After investigating $t$-ratios (used with ACAT level factors) and beta-weights (used for the sample population), we determined that the relationships were not significantly co-linear enough to substantiate causation. Consequently, there was no merit in running any further regression that analyzed the factors as predictors. However, the researchers conducted another set of tests by modulating certain respondent demographics and holding $\bar{x}$ constant.

Factor Plotting—Modulating ACAT Levels

Figure 5 shows how the factor rankings changed after isolating ACAT levels.

ACAT I

Funding and requirements factors ($F^{18}$, $F^{19}$, $F^{23}$, and $F^{26}$) previously ranked above $\bar{x}$ dropped below $\bar{x}$, while Contractor Proposal Delays ($F^{5}$) rose markedly to become the highest impact factor. Component Directed POM Adjustment ($F^{17}$) made a noticeable shift to the top nine factors (or one standard deviation above the mean).
ACAT II

Fifteen of the factors previously ranked above $\bar{x}$ dropped below $\bar{x}$ (leaving only $F^1$, $F^2$, $F^3$, and $F^{17}$). Four of the factors that fell below $\bar{x}$ included contracting-related factors ($F^4$, $F^6$, $F^8$, and $F^9$).

ACAT III

Six of the factors ($F^{16}$, $F^{18}$, $F^{19}$, $F^{21}$, $F^{23}$, and $F^{24}$) previously ranked above $\bar{x}$ dropped below $\bar{x}$. Shortages of Personnel ($F^{29}$, $F^{30}$, $F^{39}$, and $F^{51}$) and Redirection of Contractor Efforts ($F^{37}$) became more dominating issues for the respondents. Changes in User Priorities ($F^{15}$), Changes in Stakeholder Requirements ($F^{13}$), and Funding Loss from Reprograming Actions due to Higher Priority Requirements ($F^{26}$) all moved significantly above $\bar{x}$.

This more detailed differentiation, as found in the scatter plots, gives additional insight into ACAT-specific areas through a more granular view of the factors that would benefit from a more focused investigation. In some cases, reducing frequency of occurrence or perhaps instituting more early warning metrics could have a marked effect in reducing any adverse impacts.

<table>
<thead>
<tr>
<th>ACAT I</th>
<th>ACAT II</th>
<th>ACAT III</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="chart.png" alt="Chart" /></td>
<td><img src="chart.png" alt="Chart" /></td>
<td><img src="chart.png" alt="Chart" /></td>
</tr>
</tbody>
</table>

**Figure 5. Factor Ratings $\geq \bar{x}$ by ACAT Level**

**Factor Plotting—Modulating Respondent Groups**

Figure 6 shows how the factor rankings changed after isolating the Respondent Groups.

**Program Office**

Six factors dropped below $\bar{x}$: Awaiting Reprogramming Action ($F^{18}$), Changes in User Priorities ($F^{15}$), Program Delays from Prerequisite Events ($F^{21}$), Unplanned Congressional adds to PB Request ($F^{23}$), Use of Undedefinitized Contract Action Delays ($F^{24}$), and Loss of
Funding through Reprogramming Action to higher priority Requirements to PEO Portfolio (F26). No factors fell below $\bar{x}$.

\textit{PEO}

Use of Undefinitized Contract Action Delays (F24) fell below $\bar{x}$ while four factors rose above $\bar{x}$: Shortage of Cost Estimators (F26), Shortage of Business and Finance Management Personnel (F20), Component Comptroller withheld (F30), and Insufficiently Planned OCO Funding (F50).

\textit{Senior OSD Staff}

Awaiting Reprogramming Action (F18) fell below $\bar{x}$ while 13 factors rose above $\bar{x}$.

For PEO and senior OSD staff, personnel shortages (F29, F30, F20, F40) became more dominant, while awaiting reprogramming action (F18) became less dominant for program office and senior OSD staff personnel. Of the three grouping in this particular case, nowhere were there more factor increases than for senior OSD staff personnel. The rise in factors F34, F43, and F59 seemed intuitive since senior staff may see first-hand the longer time it takes for programs to react to changes in their plans. However, it was very interesting to see what senior OSD staff felt represented the major impediments to meeting the OSD’s Obligation and Expenditures rate goals that program office personnel did not, especially shortage of personnel and contract-specific factors like F45 and F37. This wide perception disparity deserves a more intensive understanding since it could be creating false perceptions that could lead to misrepresented positions and even unsubstantiated decisions.

\textbf{Figure 6. Factor Ratings $\geq \bar{x}$ by Grouped Respondent Position}
Respondent Comments Regarding the Factors

The respondents were also asked several open-ended questions about the use of metrics they found that helped them better meet OSD goals, as well as any process improvements they would recommend. They stated that the metrics that made a difference for them included “real-time monitoring, frequent reviews, tight coupling to contractor actions and milestones, and realistic spend plans with inch stones.” As for necessary improvements to current processes, the respondents recommended the inclusion of a CRA duration variable that readjusted expectations, establish more realistic program goals, ensure more funding stability, reduce bureaucratic obstacles and streamline more outdated processes, forge greater cooperation between government and industry, and synchronize disparate accounting systems used in obligation/expenditure reporting.

The respondents provided a number of qualitative comments that reinforced the quantitative data, especially for the factors above $\geq \bar{x}$ that were causing obligation rate interference:

- **$X_1$** - Personnel, Tools & Training
  - “Takes too long to get Acquisition Strategies and Acquisition Plans written and approved”
  - “Personnel do not have experience with the subject matter”

- **$X_2$** - Contracting Activities
  - “Inadequate proposals, protracted negotiations, lengthy audits, and lengthy pre-award processes”

- **$X_3$** - Requirements Stability
  - “Had to defer/re-prioritize requirements execution and carry forward funding to cover cutbacks/shortfall.”
  - “Changes in requirements precipitated by other stakeholders' actions and ill-defined requirements”
  - “User leadership routinely changes requirement & priorities”

- **$X_5$** - Business Ops
  - “MIPR billing process can delay expenditures from 90 to 120 days.
  - “Delays in negotiating best deal for govt and sometimes delays in getting acceptable proposals”

- **$X_6$** - Senior Level Execution Reviews
  - “Extensive reviews, too long to get decision briefs through oversight layers—not always value added”
  - “Multiple instances where milestone documentation took upwards of 9 months to a year to get approved”

- **$X_7$** - Funding Realities
  - “…the problem isn't unrealistic or overly optimistic spend plans as much as it's not knowing when funds will be appropriated and how much will be apportioned by the executing organization.”

---

Figure 7. Sampling of Respondent Comments
Comparison With a Similar Study

After this study was completed, the OUSD(AT&L), ODCAPE, and OSD(C) sponsored a related effort with the Institute for Defense Analyses (IDA) entitled *Implications of DoD Funds Execution Policy for Acquisition Program Management* (Conley et al., 2014). IDA was asked to increase the current understanding of the extent and causes of under-execution and suggest changes to improve outcomes. They took a two-fold approach: (1) Examine trends in the ability of the DoD to execute appropriated funds, and (2) conduct an in-depth investigation into the causes of funds under-execution for selected programs and the effects on those programs of associated financial management practices. After drawing insights from an in-depth investigation of 25 individual programs during face-to-face interviews, they categorized their causal factors along five areas:

- Contracting issues (i.e., personnel shortages and inexperience, award protests, peer reviews of contracting process documentation, and negotiation delays)
- Congressional actions (i.e., additions and reductions to requested funding, continuing Resolutions [CRs], and sequestration)
- Management actions and program events (i.e., changes to requirements, contract type, schedule, responses to operational needs, technical and testing problems, and slow contractor billing)
- Policy choices (i.e., use of execution benchmarks and withholding funding by services under CRs)
- Program office personnel shortages and experience levels

Figure 8 shows how the study results are very similar, although IDA did not measure “frequency.”
Figure 8. Comparison of DAU and IDA Studies

- **Contracting issues**
  - Personnel shortages and inexperience
  - Award protests
  - Peer reviews of contracting process documentation
  - Negotiation delays

- **Congressional actions**
  - Additions and reductions to requested funding
  - Continuing resolutions (CRs)
  - Sequestration

- **Management actions and program events**
  - Changes to requirements, contract type, schedule, responses to operational needs, technical and testing problems, and slow contractor billing
  - Policy choices
  - Use of execution benchmarks
  - Withholding funding by Services under CRs

- **Program office personnel shortages and experience levels**
Recommendations

What next? Based on the research findings of this study, there are a number of impact factors above \( \bar{x} \) that if sufficiently addressed could help lower the barriers to the attainment of the OSD’s Obligation and Expenditure rate goals. Hence, we offer the following recommendations:

- Institute an Obligation and Expenditure baseline adjustment for programs affected by any funding delay or limitation (especially CRA), then measure a program’s progress to that revised adjustment.
- More thoroughly review the entire contracting action value chain. Look closely at efficiency opportunities along the review and decision cycle continuum, especially from the time an RFP is developed to the time a contract is let. Set reasonable time thresholds with triggers that afford more proactive measures by PMs and confirm productivity.
- Establish a recurring communication forum among key stakeholders, especially PMs and the OSD, to dialogue more frequently and eliminate perception gaps that could be creating counterproductive actions and misconceptions.
- Track requirement changes throughout a program’s life and look more strategically at the effects on program execution and accompanying Acquisition Program Baselines (APBs). Despite ACAT levels, there is an obvious ripple effect that is associated with any substantive change in program content across a program’s life that should be codified more comprehensively. However, there are also issues associated with different ACAT levels which must be noted.
- Review the program review cycle and streamline wherever possible. Checks and balances within the DoD’s acquisition community are a vital constituent component of program execution, but every review should have a distinctive purpose, exit criteria, and associated suspense date that is just as material and credible.
- Build and maintain realistic spend plans, measure against them, account for contingencies, and make adjustments with required frequency due to real world realities. Collaborate with senior leadership early enough about required adjustments to avoid more draconian measures later.
- Validate the key personnel shortage areas and recognize the time it takes to rebuild those experience levels.
- Nurture experience in key functional areas with strong catalysts such as disciplined on-the-job training (OJT) programs, mentoring, and guidance. With the recent surge of contracting specialist interns, their progress as a group should be measured more carefully.
- Evaluate the real effects of reprogramming action or realignment of future budget decisions before any corrective action is taken.
- Conduct a wholesale review of the program execution metrics currently in place and determine their usefulness and effectiveness. What are they actually measuring? Consolidate whenever practical and eliminate the ones that have outlived their usefulness.
Encourage innovation and avoid the “bookkeeping process,” as RAND Corporation (2009) found in a recent study could be limiting improvements championed by PMs.

Summary

This research exposed a number of challenges that could easily be mitigated by more frequent communication and especially a better appreciation of stakeholder management. There are so many stakeholders involved in the acquisition process. No stakeholder should be dismissed without a more intensive assessment of their (potential) contribution. Sometimes, either their voice is not heard or their concern not appropriately considered. Next, having a program management strategy that can help leaders react to funding reductions is also critical. A wide variety of financial tools exist that track and predict funds execution, but Spend Plans that serve as the common device to convey program execution have to be current and agile enough to demonstrate reality and common sense for whatever curve balls come their way. Finally, as baby boomers start to retire at a more aggressive rate, experience will matter even more. An OJT program that nurtures experience and leadership development as well as demands critical thinking is just what the acquisition community should expect.

References

Contracting Processes, Internal Controls, and Procurement Fraud: A Knowledge Assessment

Juanita M. Rendon—is a CPA and instructor at the Naval Postgraduate School (NPS). Rendon teaches auditing, finance, and accounting courses at NPS. She is a former IRS revenue agent. She is a member of the American Institute of Certified Public Accountants and the Association of Certified Fraud Examiners. She has published articles in the Journal of Contract Management and book chapters in Cost Estimating and Contract Pricing (2008) and Contract Administration (2009) by Gregory A. Garrett. She has presented at the National Contract Management Association (NCMA) World Congress conferences as well as at NCMA's National Education Seminars, and was the recipient of the 2008 NCMA Award for Excellence in Contract Management Research and Writing. [jmrendon@nps.edu]

Rene G. Rendon—is an associate professor at NPS, where he teaches defense acquisition courses. Prior to his appointment at NPS, he served for more than 22 years as an acquisition and contracting officer in the United States Air Force, retiring at the rank of lieutenant colonel. His Air Force career included assignments as a contracting officer for the Peacekeeper ICBM, Maverick Missile, and the F-22 Raptor. He was also the director of contracting for the Air Force’s Space-Based Infrared satellite program and the Evolved Expendable Launch Vehicle rocket program. [rgrendon@nps.edu]

Abstract

In fiscal year 2014, the U.S. federal government obligated over $443 billion in contracts. Within the Department of Defense (DoD), over $283 billion were obligated in the execution of 1.3 million contract actions (USA Spending, 2015). Despite the critical importance of the DoD’s contract management function, both the Government Accountability Office (GAO) and the DoD Inspector General (DoD IG) have reported problems in achieving successful procurement outcomes. The lack of trained personnel, capable processes, and effective internal controls result in the DoD having a higher level of vulnerability for procurement fraud (Rendon & Rendon, in press). Contracting officers, because of their pivotal position in the procurement process and their interface with both government officials and industry, are in a unique position to be on the front line for deterring and identifying procurement fraud. However, in order to be effective procurement fraud fighters, contracting officers must be knowledgeable of both contracting processes and internal controls. The purpose of this research was to assess DoD contracting officers’ knowledge of the DoD’s contract management processes and related internal controls. Our research findings indicated contracting officers may have a possible knowledge deficiency in the areas of procurement internal controls and procurement fraud schemes. Based on the implications of these findings, recommendations are made to the assessed agency and the DoD.

Background

In fiscal year 2014, the U.S. federal government obligated over $443 billion in contracts. Within the Department of Defense (DoD), over $283 billion were obligated in the execution of 1.3 million contract actions (USA Spending, 2015). These contract actions were for the procurement of weapon systems, supplies, and services. Despite the critical importance of the DoD’s contract management function, both the Government Accountability Office (GAO) and the DoD Inspector General (DoD IG) have reported problems in achieving successful procurement outcomes. The GAO has identified the lack of trained contracting personnel and the use of ill-suited contracting arrangements as some of the problems in DoD contract management (GAO, 2013). The DoD IG has identified numerous deficiencies in contract management processes as well as weakness in contract management internal controls (DoD, 2009, 2014). The lack of trained personnel, capable processes, and effective internal controls result in the DoD having a higher level of vulnerability for procurement fraud (Rendon & Rendon, in press). Past incidents of procurement fraud can be traced to
incompetent personnel, incapable contracting processes, or ineffective internal controls (Tan, 2013). Contracting officers, because of their pivotal position in the procurement process and their interface with both government officials and industry, are in a unique position to be on the front line for deterring and identifying procurement fraud. However, in order to be effective procurement fraud fighters, contracting officers must be knowledgeable of both contracting processes and internal controls.

The purpose of this research was to assess DoD contracting officers’ knowledge of the DoD’s contract management processes and related internal controls. Specifically, we focus on the following research questions:

1. What is the contracting workforce’s knowledge level of procurement fraud schemes as related to contract management processes, internal control components, and procurement fraud scheme categories?
2. What is the contracting workforce’s perception of procurement fraud as related to the contract management processes, internal control components, and procurement fraud scheme categories?

Literature Review

This section provides a brief literature review that serves as a foundation for our empirical study. We review academic journal articles, government reports, and previous research studies in the areas of auditability, contract management processes, and internal control components.

Auditability in Public Organizations

The literature on auditability reflects an organization’s transformation in its governance and knowledge management capabilities. Power (2007) states, “Auditability is a condition of possibility of all inspection and auditing practices and also a mode of organizational transformation” (p. 14). This organizational transformation occurs when data collection practices and documentation systems are established, thus allowing organizations to be audited. This is distinct from organizations conducting actual audits or inspections. Audibility requires organizations to establish and actively manage an institutionally acceptable knowledge management system supporting its governance of processes and practices (Power, 1996). Organizations’ increased concern for risk management has resulted in an accompanying emphasis on auditability of internal controls (Power, 2007).

Auditibility is also reflected in an organization’s governance structure for the management of procurement activities. The literature includes research supporting the importance of competent personnel (in terms of education, training, and experience) and competent organizations (in terms of capable processes) for ensuring the success of procurement projects (Frame, 1999) and the need for a renewed emphasis on strong internal controls as a response to the increase in procurement fraud incidents (Rollins & Lanza, 2005). Finally, the literature also includes past research supporting governance structures in public sector agencies and the role procurement projects play in ensuring accountability, transparency, compliance, and consistency in delivery, value for money, and stakeholder engagement (Crawford & Helm, 2009).

Contract Management Processes

A common focus of theoretical research on contract management is the use of an agency theory lens (Jensen & Meckling, 1976). Agency theory, specifically the buyer–seller problem, is used to analyze the process of structuring contract agreements between the buyer (principal) and the seller (agent) for the performance of a service or the delivery of a
product (Eisenhardt, 1989). Because of conflicting goals and asymmetrical information between the buyer and the seller, as well as the opportunistic behavior of both parties, contracts are used to govern the buyer and seller relationship. Structuring contracts using product, exchange, and governance rules allows both parties to align contract goals at the lowest cost. Product rules are used to establish product specifications or service requirements, and exchange rules specify the parties’ rights and obligations, as well as contract period of performance, delivery schedule, and method of contractor payment. Governance rules are used to reward and sanction cooperative or defective behavior of both parties through the use of performance incentive or penalty clauses (Brown, Potoski, & Van Slyke, 2013).

The structuring of contracts follows the generally-accepted contract management phases of procurement planning, solicitation planning, solicitation, source selection, contract administration, and contract closeout (Rendon & Snider, 2008).

Procurement planning involves the process of identifying which business needs can be best met by procuring products or services outside the organization. This process involves determining whether to procure, how to procure, how much to procure, and when to procure. The solicitation planning process involves preparing the documents needed to support the solicitation. This process involves documenting program requirements and identifying potential sources. Solicitation is the process of obtaining information (proposals) from the sellers on how project needs can be met. The source selection process includes evaluating proposals and conducting contract negotiations with the seller in an attempt to come to agreement on all aspects of the contract—including cost, schedule, performance, terms and conditions, and anything else related to the contracted effort. Contract administration is the process of ensuring that each party’s performance meets the contractual requirements. The activities involved in contract administration will depend on the contract statement of work, contract type, and contract performance period. Contract closeout/termination is the process of verifying that all administrative matters are concluded on a contract that is otherwise physically complete. A government contract can end in one of three ways. First, the contract can be successfully completed, allowed to run its full period of performance, and then closed out. Second, the contract can be terminated for the convenience of the government. Finally, the contract can be terminated for default. Regardless of how the contract ends, all contracts must be closed out.

These contract management processes will only be as capable and effective as the internal controls used by the acquisition agency to manage and oversee those processes. The next section provides an overview of internal controls as applied to acquisition agencies.

Internal Controls

As discussed previously, having capable contracting management processes helps organizations become auditable. In addition to capable contracting processes, effective internal controls are important in order for organizations to become auditable.

In 1992, the Committee of Sponsoring Organizations (COSO), composed of the AICPA, IIA, FEI, and the AAA, established the Internal Control Integrated Framework, which includes five internal control components (COSO, 2013). In May 2013, COSO (2013) updated its internal control integrated framework, which now includes 17 principles within the five components of internal control. In September 2014, the GAO (2014) updated its Standards for Internal Control for the Federal Government (Green Book). The five components of the framework are discussed in the following sections (COSO, 2013).
Control Environment
The control environment component of the framework entails the tone at the top. Management’s integrity and ethical behavior sets the tone for the organization (COSO, 2013). A weak control environment can open the door to fraud, waste, and abuse (GAO, 2006). The effects of waste and abuse can be just as damaging as fraud to any organization in terms of loss of dollars, time, and personnel. In the case of government organizations, public trust could be compromised and public funds could be lost when the control environment is weak.

Risk Assessment
The risk assessment component of the framework calls for management to discuss what could go wrong within the organization and how to best mitigate any potential risks, including fraud risks (COSO, 2013).

Control Activities
The control activities component of the framework encompasses all of the control procedures that have been determined to be needed to make sure that the organization meets its goals and objectives (COSO, 2013). One example of a control activity is segregation of duties or separation of duties. No one person should be in charge of all the procedures within a process. This could lead to opportunities for unscrupulous people to commit fraud.

Information and Communication
The information and communication component of the framework includes the accounting system and the methods of internal and external communications within an organization (COSO, 2013).

Monitoring Activities
The monitoring activities component of the framework helps ensure that the controls in place are being followed and are meeting the organizational goals set by management (COSO, 2013). If any controls need to be updated, changed, removed, or added, management can determine the best way to proceed. Monitoring activities is important to help ensure a continuous process of planning, implementing, reviewing, and adjusting controls.

The previously discussed internal control components are integral to ensuring auditability in the acquisition agency’s procurement and contracting processes. Any material weaknesses in the agency’s internal control components can increase its vulnerability to procurement fraud. Although there are many different types of procurement fraud incidents, the majority of procurement fraud can be categorized in the following procurement fraud schemes.

Procurement Fraud Scheme Categories
Reducing contract fraud, waste, and abuse should be the goal of any government organization, especially the DoD (GAO, 2006). Ineffective internal controls leave government organizations vulnerable to contract fraud, waste, and abuse, as shown in the many incidents of procurement fraudulent activity within the DoD and the federal government (Tan, 2013). While there are numerous kinds of procurement fraud, they can be classified into six categories, which include collusion, bid rigging, conflict of interest, and billing/cost/pricing schemes, fraudulent purchases, and fraudulent representation. These fraud scheme categories are discussed next.
**Collusion**

Specific types of fraud schemes included within the collusion category of fraud schemes are kickbacks, bribery, and deliberate split purchases. Kickbacks involve government officials receiving something of value such as money from a contractor for personal gain in exchange for providing a favor such as submitting false invoices. Bribery involves influencing someone’s judgment in order to obtain favor, such as bribing a contracting officer in order to be awarded additional contracts. Split purchases are often seen in the government purchase card program where purchases that generally would not meet the micro-purchase threshold of $3,000 are deliberately split into two or more purchase transactions to circumvent the contracting rules and regulations.

**Bid Rigging**

Bid rigging involves exploitation of the bidding process by falsifying information such as price competition, agency needs, and contract specifications with the intent to circumvent the standard bidding process. Government bid specifications purposely leaked to favored offerors feeds the bid rigging fraud scheme and creates an unfair advantage to others seeking government contracts (Wells, 2008).

**Conflict of Interest**

Conflicts of interest create problems for government officials who are in a position to make decisions that could be seen as not being in the best interest of the government. For example, a contracting officer who is reviewing a contract bid from a company in which he or she owns stock would be construed as a major conflict of interest.

**Billing/Cost/Pricing Schemes**

Billing, cost, or pricing type fraud schemes involve a misrepresentation of financial information as well as intentionally mischarging. For example, government losses are sometimes due to a contract’s labor cost mischarging, such as padding employee timecards and charging the government for the extra hours that were not worked by employees. Defective pricing, change order abuse, and comingling of contracts are also examples of cost and pricing schemes.

**Fraudulent Purchases**

Fraudulent purchases involve purchases made which are beyond the government requirements with the intent to defraud the government. An example of fraudulent purchases is when a purchase cardholder purchases electronic equipment for personal use or with the intent to sell the equipment on e-bay for personal gain.

**Fraudulent Representation**

Fraudulent representation involves falsely and intentionally misrepresenting goods and services. Product substitution is an example of fraudulent representation, as it usually involves intentionally providing defective or used parts instead of non-defective or new parts as required by a contract. The following section discusses the conceptual framework of this research.

**Conceptual Framework**

The auditability literature identifies the importance of competent procurement personnel, capable contract management processes, and effective procurement internal controls. These major facets of auditability are reflected in Figure 1, which presents our conceptual framework.
Our conceptual model shows that auditability in procurement agencies requires organizations to have competent people, capable processes, and effective internal controls. The acquisition workforce needs competent people that are educated, trained, and experienced in the complexities of government contracting. Past research has shown that many participants in the acquisition of services, especially at the installation/base level, are not members of the acquisition workforce. Thus, they may not be receiving the required education and training needed to perform their acquisition functions (GAO, 2002, 2011).

Acquisition organizations also need capable contract management processes. Process capability is measured in terms of processes that are fully-established, institutionalized, mandated, integrated with other organizational processes, periodically measured, and continuously improved. Past research using the Contract Management Maturity Model (CMMM) has shown that some acquisition organizations have less-than-capable contracting processes. These contracting processes lack process strength, management support, process measurement, and process improvement (Rendon, 2008, 2009, 2010, 2011).

Finally, acquisition organizations also need effective internal controls. Effective internal controls refers to the objective of enforcing internal control policies to ensure compliance with laws and regulations, monitoring procedures to assess enforcement, and reporting material weaknesses. The DoD IG (2009, 2014) reports that many of the procurement deficiencies they identified are the result of material internal control weaknesses in the procurement processes. Both GAO and DoD IG reports have indicated an ever-increasing concern about weak internal controls within the DoD’s acquisition agencies. These types of internal control weaknesses increase the government’s risk of jeopardizing the value for the public dollars spent on supplies and services (DoD IG, 2009, 2014).

The lack of competent personnel, capable processes, and effective internal controls may be resulting in incidents of procurement fraud within the DoD and throughout the defense supply chain (Tan, 2013). Government contracting officers, because of their pivotal position interfacing with industry, should be the first line of defense for identifying procurement fraud red flags. However, identifying procurement fraud red flags require contracting officers to be knowledgeable of procurement internal controls. Thus, the purpose of this research is to assess DoD contracting officers’ knowledge of the DoD’s contract...
management processes and related internal controls. We conduct our research through the use of a web-based assessment tool. Our research methodology is discussed next.

**Research Methodology**

We conducted this research by first developing and testing a web-based assessment tool that can be used to assess the contracting officers’ knowledge level of internal controls and procurement fraud and to assess their perceptions of procurement fraud within their organizations. With the assistance of our MBA student (Chang, 2013), we developed a web-based assessment tool consisting of 26 knowledge-based questions pertaining to contracting processes, internal control components, and procurement fraud schemes. These knowledge-based questions were developed using government procurement fraud sources such as the U.S. Agency for International Development (USAID, n.d.) and the Office of the Inspector General, U.S. General Services Administration (GSA) procurement fraud handbook (GSA OIG, 2012). The assessment tool also included 12 organization-based items related to the contracting officers’ perceptions of internal controls within their organizations. These items were designed to determine if any aspects of the organizations’ internal structure, processes, or culture made the organization more susceptible to fraudulent activity. These items were also designed to assess the contracting officers' perceptions of their organizations regarding fraud incidents. The organization-based items were adopted and modified from the Internal Control Survey developed by the New York State Internal Control Association (NYSICA, 2006).

The survey respondents were contracting officers assigned to the Army Mission Installation Contracting Command (MICC). This Army contracting agency is responsible for supporting the U.S. Army installations by contracting for office supplies, equipment, support services, and minor construction. In 2012, the MICC managed over $6.4 billion in contracts using a variety of contract mechanisms, ranging from the government purchase card to complex services contracts (Chang, 2013).

**Research Findings**

The assessment tool was deployed in early April 2013 to a total eligible population of 1350 contract management professionals. The assessment tool was initiated by 146 respondents, and was completed by 99 respondents, resulting in a response rate of 7% (Chang, 2013).

**Analysis of Knowledge Assessment Findings**

The average score on the knowledge portion of the assessment tool was 63% correct of the 26 questions. There were minimal differences in average scores between civilian and military contract management professionals. However, there were some differences in average scores by experience and DAWIA levels. As contracting experience and DAWIA level increases, so does the average score on the knowledge assessment. Although warranted contracting officers scored higher than non-warranted contracts specialists, there is less difference in average scores between non-warranted and warranted contract management professionals.

Each knowledge assessment item was related to contract management processes, internal control components, and procurement fraud schemes. Figures 2 through 4 reflect the average score based on each of these areas.

As can be seen in these figures, there is variation in the average knowledge assessment score among these three areas. From the perspective of the contract management process, assessment knowledge items related to the procurement planning...
process had the highest average score, compared to items related to contract closeout, which had the lowest score. From the perspective of the internal control components, assessment knowledge items related to the risk assessment component had the highest average score, compared to items related to information and communication, which had the lowest score. From the perspective of procurement fraud schemes, assessment knowledge items related to bid rigging scheme had the highest average score, compared to items related to billing/cost/pricing schemes, which had the lowest score.

Figure 2. Average Score by Contract Management Process (Chang, 2013)
Analysis of Organization Perception Findings

The assessment tool also included items related to perceptions of the organization’s vulnerability to procurement fraud and perceptions of the organization’s internal controls.
Three of these questions asked about perceptions of the organizations’ vulnerability to procurement fraud in terms of contract management process, internal control component, and procurement fraud schemes. Figures 5–7 reflect the responses to these items.

As can be seen in these figures, when asked which contract management process is most vulnerable to fraud in their organization, those surveyed selected the procurement planning process the most often (20%) and contract closeout was selected the least often (0%). Approximately 11% responded that they did not know, and 34% of the respondents stated they did not suspect fraud.

When asked which internal control component is most vulnerable to fraud in their organization, the information and communication component was selected the most often (13%) and control environment was selected the least often (4%). Approximately 17% responded that they did not know, and 38% of the respondents stated they did not suspect fraud.

When asked to which procurement fraud scheme they perceived their organization was most susceptible, those surveyed selected conflict of interest the most often (13%), and fraudulent representation was selected the least often (0%). Approximately 13% responded that they did not know, and 53% of the respondents stated they did not suspect fraud.

![Figure 5. Percent of Responses to Contract Management Phase Item (Chang, 2013)](image)
Figure 6. Percent of Responses to Internal Control Component Item
(Chang, 2013)

Figure 7. Percent of Responses to Procurement Fraud Scheme Item
(Chang, 2013)
Implications of Findings

The results of both the knowledge assessment and the organization perception assessment have some interesting implications. The average score on the knowledge assessment varied by contract management process, internal control component, and procurement fraud scheme. The contracting officers’ average score on the knowledge assessment (63%) reflects a possible knowledge deficiency in procurement internal controls. Using traditional college grading protocol, this score would be converted to a grade of D. This finding, along with the average response to the organization perception item, “I have adequate knowledge of contracting fraud schemes to perform my duties” (see appendix, Item 6) of 3.9, suggests that perhaps the contracting officers are overly optimistic in self-assessing their knowledge of procurement fraud schemes.

Additionally, a significant percentage of the respondents indicated that “I do not suspect fraud” in relation to the organization’s contracting processes (34%), internal control components (38%), and procurement fraud scheme susceptibility (53%). These findings, along with the low scoring knowledge assessment, may indicate that although the majority of contracting officers do not suspect fraud in their organizations, they also do not have a sufficient working knowledge of procurement fraud. The contracting officers’ limited knowledge of procurement fraud and their perception that their organization is not susceptible to fraud may reveal that the organization could in fact be vulnerable to some form of procurement fraud.

Recommendations

The results of the knowledge-based assessment indicated that, although the average score was 63%, the contracting officers’ knowledge of contracting processes, internal controls, and procurement fraud schemes increased as years of experience and DAWIA certification level increases. Recent research shows that the DAWIA required courses for contracting certification do not include a mandatory fraud training or awareness course (Castillo & Flannigan, 2014). Our first recommendation is for the Defense Acquisition University (DAU) to incorporate coverage of internal controls and procurement fraud schemes in the mandatory contracting curriculum. Our final recommendation is to further explore the organization’s information and communication internal control component. This was the internal control component with the lowest score on the knowledge assessment, as well as the component chosen as most vulnerable to procurement fraud in the organization. Perhaps this organization should apply additional emphasis—for example, training and visibility—in this internal control component. This may increase the workforce’s knowledge level of this aspect of their organization’s internal controls and decrease their perception of this area of fraud vulnerability.

Conclusion

Auditability theory states that an organization must have competent personnel, capable processes, and effective internal controls to ensure proper organizational governance. The lack of competent personnel, capable processes, and effective internal controls may result in organizations being more vulnerable to fraud. The purpose of this research was to assess DoD contracting officers’ knowledge levels of procurement fraud schemes as related to contract management processes, internal control components, and procurement fraud scheme categories. Our research findings indicated contracting officers may have a possible knowledge deficiency in the area of procurement internal controls. Additionally, our findings indicate that perhaps contracting officers are overly optimistic in self-assessing their knowledge of procurement fraud schemes. Finally, we also conclude
that the contracting officers’ limited knowledge of procurement fraud and their perception
that their organization is not susceptible to fraud may reveal that the organization could in
fact be vulnerable to some form of procurement fraud.

Based on these findings, we recommend that the DAU incorporate coverage of
internal controls and procurement fraud schemes in the mandatory contracting curriculum.
We also recommend that the assessed organization apply additional emphasis on its
information and communication internal control component to increase the workforce’s
knowledge level of this aspect of their organization’s internal controls and decrease their
perception of this area’s vulnerability to fraud.

References


5. I know who to report to if I saw or suspected fraudulent activities.

<table>
<thead>
<tr>
<th>Response</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Agree</td>
<td>4.31</td>
</tr>
<tr>
<td>Agree</td>
<td>4.01</td>
</tr>
<tr>
<td>Neither Agree nor Disagree</td>
<td>3.00</td>
</tr>
<tr>
<td>Disagree</td>
<td>2.00</td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>1.00</td>
</tr>
</tbody>
</table>

I Don’t Know: 3

6. I have adequate knowledge of contracting fraud schemes to perform my duties.

<table>
<thead>
<tr>
<th>Response</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Agree</td>
<td>3.90</td>
</tr>
<tr>
<td>Agree</td>
<td>3.60</td>
</tr>
<tr>
<td>Neither Agree nor Disagree</td>
<td>2.60</td>
</tr>
<tr>
<td>Disagree</td>
<td>1.60</td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>0.60</td>
</tr>
</tbody>
</table>

I Don’t Know: 2

7. Instances of reported suspected fraudulent or suspicious activity have been adequately investigated by my organization.

<table>
<thead>
<tr>
<th>Response</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Agree</td>
<td>3.84</td>
</tr>
<tr>
<td>Agree</td>
<td>3.54</td>
</tr>
<tr>
<td>Neither Agree nor Disagree</td>
<td>2.54</td>
</tr>
<tr>
<td>Disagree</td>
<td>1.54</td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>0.54</td>
</tr>
</tbody>
</table>

I Don’t Know: 37

8. Employees in my organization who are found to have participated in fraudulent activities will be subject to appropriate consequences.

<table>
<thead>
<tr>
<th>Response</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Agree</td>
<td>4.08</td>
</tr>
<tr>
<td>Agree</td>
<td>3.78</td>
</tr>
<tr>
<td>Neither Agree nor Disagree</td>
<td>2.78</td>
</tr>
<tr>
<td>Disagree</td>
<td>1.78</td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>0.78</td>
</tr>
</tbody>
</table>

I Don’t Know: 23
9. My organization places sufficient emphasis on the importance of integrity, ethical conduct, fairness and honesty in their dealings with employees, vendors, and other organizations.

4.34

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I Don't Know: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Department of Defense Management of Unobligated Funds for Acquisition Programs

William Fast—has been a senior lecturer in acquisition in the Graduate School of Business and Public Policy at the Naval Postgraduate School since September 2010. Prior to that, he taught acquisition and business, cost estimating, and financial management courses at the Defense Acquisition University. Fast is credentialed by the American Society of Comptrollers (ASMC) as a Certified Defense Financial Manager with Acquisition Specialty (CDFM-A). [wrfast@nps.edu]

Abstract

On September 10, 2012, the Under Secretary of Defense (Comptroller) and the Under Secretary of Defense for Acquisition, Technology, and Logistics signed a joint memorandum calling for a change in the management of unobligated funds for acquisition programs. The memo was an attempt to stop the rush to obligate funds by the end of the current fiscal year in order to avoid reduced allocations in the next fiscal year. The memo identified six tenets for adoption throughout the Department of Defense (DoD). Included are tenets to reward acquisition program managers for returning unobligated funds and not to reduce future year program budgets just because current year obligation rates are lower than established benchmarks. The 2014 Annual Report on the Performance of the Defense Acquisition System (OUSD[AT&L], 2014) indicated that the DoD was continuing the effort to change the acquisition culture from focusing on obligation rates. This research analyzes financial obligation rates for acquisition programs and acquisition program contract awards made in the last quarter of the fiscal year to determine if this policy memo has resulted in any change in behavior toward year-end spending. Research finds that the policy memo has not yet had an effect on behavior and that it has not been incorporated into DoD policy regulations.

Background

Unobligated balances are amounts of budget authority available to acquisition program managers (PMs) that have not yet been obligated on contracts. Each year, PMs must plan for obligating a high percentage of their available budget authority or risk losing the unobligated balance to a higher priority program. The process for repurposing budget authority from one program to another or from one appropriation account to another is termed reprogramming, and is done by the component comptroller and the Under Secretary of Defense (Comptroller). There are three primary means of reprogramming that could be used to repurpose unobligated balances: below threshold, congressional notification letter, and prior approval.

Budget authority is provided to the Department of Defense (DoD) by three annual appropriation acts: the Defense Appropriation Act; the Military Construction, Veterans Affairs and Other Related Agencies Appropriation Act; and the Energy and Water Development Appropriation Act. Once these acts are signed by the President, the DoD can request apportionment of the budget authority from the Office of Management and Budget (OMB). The OMB usually apportions the annual amount of budget authority for the procurement appropriation accounts and either annual or quarterly amounts for Research, Development, Test, and Evaluation (RDT&E) appropriation accounts. However, the president may temporarily impound budget authority for programs with technical problems and may request rescission of budget authority for programs that have been canceled or changed significantly. In addition, the OMB may reduce the apportioned budget authority to pay for a program that the Congress authorized, but did not appropriate any budget authority for in one of the appropriations acts. For example, each year the Congress authorizes the Small
Business Innovative Research (SBIR) program, but directs the DoD to fund the SBIR program with a general reduction in budget authority from all RDT&E-funded programs.

In anticipation of the apportionment of budget authority, all PMs are to submit obligation and expenditure plans (also termed spending plans) to their component comptroller. These plans serve as a month-by-month forecast, through the entire fiscal year, of when budget authority will be obligated on contract and when the obligated amounts will be liquidated by expenditures made to pay defense contractors. It is important for the PM to prepare realistic spending plans. For new start programs, obligation of budget authority should not be planned for the first quarter of the fiscal year. Recent history has shown that the DoD operates in the first quarter of most fiscal years under stop-gap funding called a continuing resolution. Since the annual appropriation bills are still in work, budget authority for new starts has not yet been determined or appropriated. Therefore, obligations for new starts cannot be made while operating under a continuing resolution. PMs are also cautioned not to plan for the obligation of large amounts of budget authority in the fourth quarter of the fiscal year. This is because the appropriation acts usually restrict obligations in the last two months of the fiscal year to no more than 20% of the appropriated amount. The purpose of this constraint is to force planning and avoid rushing into contracts at the end of the fiscal year just to show that fund obligations have been made.

In the March timeframe of each fiscal year, the component comptroller usually conducts a mid-year review of the status of obligations and expenditures by program (see Figure 1).

![Figure 1. Obligation Availability Timeline for RDT&E Appropriations](image)

The purpose of this review is to identify programs that have deviated from their spending plans and have large unobligated balances of budget authority (i.e., low obligations) or large unliquidated obligations (i.e., low expenditures). About two years earlier, in the budget formulation process, the PM should have asked for RDT&E funds to cover all contractor-incurred costs in the fiscal year. This is in accordance with the incremental funding policy that applies to all RDT&E appropriation accounts (OUSD[AT&L],
Similarly, the PMs should have requested budgets to fully fund complete, militarily-useable end items in the fiscal year of contract award for those end items. This is in accordance with the full funding policy that applies to all procurement accounts, including Ship Building and Conversion, Navy, and to the Military Construction appropriation accounts (OUSD[AT&L], 2015). If budget authority is not being placed on contract in the year for which it was requested, the comptroller could decide to reprogram some or all of the unobligated amounts to a higher priority program. In Pentagon parlance, the under-obligated program could become a bill payer for a higher priority program.

If the PM successfully defends the under-obligated program at the mid-year review, the pressure is on to get the majority of the budget authority obligated by September 30, the end of the fiscal year. If the funds are not obligated by the end of the first fiscal year of availability, then the comptroller has yet another opportunity to reduce the program’s future budget request during the Budgeting phase of the Planning, Programming, Budgeting and Execution (PPBE) process. During this budget review, which occurs in the August to October timeframe, the Under Secretary of Defense (Comptroller) and OMB analysts test the Budget Estimate Submissions (BES) submitted by the components to ensure that each program’s budget request is defendable before the Congress. As a result of this PPBE process budget review, the analysts will sometimes find that a program is forward financed. That is, compared to its needs, the program has too much current and carried-over budget authority. If this is the case, the analysts will recommend that the requested future budget be reduced accordingly.

The tensions created by the mid-year and PPBE budget reviews can cause the PM to behave in different ways. First, as discussed above, the PM might decide to move quickly to get the budget authority on contract, regardless of how bad the business deal that is for the taxpayer. Obviously, this behavior is irresponsible, but logical in terms of retaining current budget authority and stabilizing future budget requests. Second, the PM might agree with the comptroller’s decision to reprogram unobligated funds to a higher priority program, provided that a promise is made to replace the budget authority removed with future funds, escalated as appropriate to account for future inflation and outlay rates. Finally, the PM might decide to fight any action by the comptroller to remove funds by sharply rebutting the action. This approach would require that a reclama be prepared, citing operational mission and supportability impacts to the user (warfighter) and business impacts that affect the execution of the program. If successful with the reclama, the PM has more time in which to negotiate a good deal before obligating the budget authority on contract.

Having observed all of these behaviors over many years, in September 2012, the Under Secretary of Defense (Comptroller), the Honorable Robert F. Hale, and the Under Secretary of Defense for Acquisition, Technology, and Logistics, the Honorable Frank Kendall III, signed a joint memorandum (copy in appendix) that outlined six new tenets for dealing with unobligated balances during budget reviews:

1. Taxpayer funds should be obligated and ultimately expended only in the taxpayers’ interest and if best value is received for the money in support of the Warfighter.
2. While they can be useful indicators, obligation rates slower than established benchmarks should not be the determinative measuring stick for program execution and must not be regarded as a failure.
3. Late obligation of funds should not be presumed to imply that the funds are not needed or that future budgets should be reduced unless there is other evidence to support that conclusion.
4. Providing savings to the organization, military service, or DOD component as early in the fiscal year as possible should be encouraged and rewarded, professionally and visibly.

5. Savings will not be reallocated at any higher DOD level than necessary to fulfill shortfalls in priority requirements.

6. Managers who release unobligated funds to higher priorities will not automatically be penalized in their next year's budget with a lower allocation and may be candidates for additional funding to offset prior year reductions. (OUSD[AT&L] & OUSD[C], 2012)

The two under secretaries asked that acquisition and financial managers throughout the DoD follow these tenets when reviewing programs not meeting established obligation rates. However, as of this writing, the policy memorandum has not been incorporated into any DoD regulations (M. Engelking, personal communication, February 5, 2015).¹

Before incorporating these tenets into DoD regulations, data-driven research should be conducted to answer three important questions:

1. What are the recent obligation rates, and how big are unobligated balances in the RDT&E and procurement accounts?
2. What is happening to unobligated balances at the end of obligation availability? Are PMs rushing into poor contracts just to prevent funds from being taken away, or to prevent reductions to future budgets?
3. What recent legislation has been introduced and/or enacted by the Congress on the subject of unobligated balances and meeting obligation benchmarks?

The balance of this paper provides data-driven research that answers these questions and can be used to guide the implementation of policy guidance on the management of unobligated balances.

**Recent Trends in Obligation Rates and Unobligated Balances**

With the rollout of the president’s budget in February of each year, the Under Secretary of Defense (Comptroller) presents a budget briefing for the news media and makes available financial summary tables. These tables contain historical and forecasted obligation rates for the RDT&E and procurement appropriation accounts. These rates, rolled to the DoD level and representing obligations at the end of the first year of obligation availability, were extracted and plotted in Figure 2.

¹ Email, dated February 5, 2015, regarding implementation of obligation rates tenets, follows: “Volume 3, Chapter 8 [DoD 7000.14-R, Financial Management Regulation] was revised and forwarded to the Office of Legal Counsel for coordination a few months ago. It was revised by an action officer who has since left OUSD(C). I am now responsible for the chapter. I reviewed my predecessor’s files related to this chapter and did not locate a copy of the memorandum you provided. I also scanned the document to see if it addressed this issue. Unfortunately, the revised version does not address it. I will request that the chapter be returned to me for revision. It will be scheduled to be published by the end of the fiscal year. Thank you for bringing this to our attention. Maryla Engelking, CPA, CGFM, MBA, Senior Staff Accountant, Office of the Under Secretary of Defense (Comptroller), Financial Improvement and Audit Readiness (FIAR), 703-571-1657, maryla.e.engelking.civ@mail.mil.”
From FY 2003 to FY 2012, RDT&E obligation rates held steady at 90%. This percentage is in accordance with historical DoD and Service/Component benchmarks for RDT&E obligations at the end of the first year of funds availability. However, overall RDT&E obligation rates dipped to 85% in FY 2013 and FY 2014. Perhaps this recent dip reflects implementation of the joint memo on obligation rate tenets, which was just discussed. However, if that were the case, one would expect to see a similar trend in the procurement accounts, which is not the case (see discussion that follows).

From FY 2003 to FY 2007, procurement obligation rates hover between 80% and 85%. This percentage is in accordance with historical DoD and Service/Component benchmarks for procurement obligations at the end of the first year of funds availability. However, procurement obligation rates begin to dip below 80% in FY 2008 and get as low as 67% and 68% in FY 2013 and FY 2014, respectively.

The financial summary tables provided with the annual president’s budget rollout briefing also contain unobligated balances, brought forward from the previous fiscal year. These balances were extracted from the tables and plotted on bar graphs for RDT&E (Figure 3) and Procurement (Figure 4). In each figure, the bottom solid bar represents the budget authority appropriated in that particular fiscal year. The shaded tip of the bar represents the carryover of unobligated budget authority from the previous year. Together, the appropriated budget authority and the carryover funds can be considered a rough approximation of the total obligation authority available in any particular year.
Figure 3. Unobligated RDT&E Appropriation Balances (Carryover) From Previous Fiscal Year
(OUSD[C], 2015b)

Figure 4. Unobligated Procurement Appropriation Balances (Carryover) From Previous Fiscal Year
(OUSD[C], 2015b)
In the DoD's four RDT&E accounts, carryover as a percentage of total obligation authority has nearly doubled from 8.8% in FY 2003 to 15.9% in FY 2014. In the DoD's procurement accounts, carryover as a percentage of total obligation authority has more than doubled from 15% in FY 2003 to 31.6% in FY 2014. Because the increases in unobligated balances began at least nine years earlier, the September 2012 joint memo on obligation rate tenets did not cause these carryover increases. Analysis of the Service/Component unobligated carryover indicates that

from analyzing the trends in obligation rates and carryover data, the answer to the first important question becomes clear. Obligations rates in both the RDT&E and procurement accounts are falling, with procurement obligation rates falling at further than RDT&E obligation rates. Also, for both RDT&E and procurement, carryover balances have more than doubled as a percentage of total obligation authority. As will be seen later, lower obligation rates and higher carryover amounts make DOD appear to be forward financed and expose future budgets to higher risks of reduction during the congressional enactment process.

**Disposition of Unobligated Balances**

At the mid-year review, or whenever a higher priority need arises, reprogramming provides some flexibility for the DoD to repurpose budget authority. Prior approval reprogramming, requiring approval of the Congressional Defense Committees before implementation, is used when budget authority changes from one appropriation to another. However, below certain dollar and percentage thresholds, budget authority can be moved from one RDT&E program element to another or from one procurement line item to another. Such below-threshold reprogramming (BTR) can be accomplished by the Service/Component and DoD comptrollers and is reported to the Congress quarterly on DD 1416 Program Reports. The Under Secretary of Defense (Comptroller) website contains only a few years of data on recent BTR actions. This data is plotted, by the four DoD RDT&E accounts, in Figures 5 and 6.

![Figure 5. Number of Below-Threshold Reprogramming Actions, RDT&E Appropriation Accounts (OUSD[C], 2014)](chart.png)
Figures 5 and 6 reveal that, at least for two years (FY 2011 and FY 2012), the number of BTR actions and the budget authority repurposed through BTR actions declined in all four of the RDT&E appropriation accounts. One might conclude from this data that the policy memo on obligation rate tenets has had an effect. Fewer BTR actions might mean that fewer acquisition programs are being tapped as bill payers for higher priorities. However, more research is needed into the reasons for the BTR actions and the reason for the decline in the number of BTR actions/amounts. Only the 2010 DD 1416 Quarterly Reports of Programs identify reason codes associated with the BTR actions. After 2010, the codes and hyperlinks to any amplifying information are missing. Help from the Office of the Under Secretary of Defense (Comptroller) is needed to provide this missing information.

Contract awards made late in the fiscal year, particularly in the fourth quarter, may also be an indicator of the effect of the policy memo on obligation rate tenets. PMs and business financial managers (BFMs) are taught not to plan for obligations of funds in the fourth quarter of any fiscal year. One reason for this is because the appropriation acts usually restrict obligations in the last two months of the fiscal year to no more than 20% of the appropriated amount. Another reason for not planning fourth quarter contract awards is that program schedules and contract awards often slip. If program schedule slippage does occur, it is easy for the comptroller to recommend that contract awards scheduled for the fourth quarter be rescheduled into the next fiscal year.

The Defense Acquisition Management Information Retrieval (DAMIR) has a database of 421 large active contracts from Major Defense Acquisition Program and Major Automated Information Systems. The award dates of these contracts are plotted in the bar chart in Figure 7.
Although PMs and BFMs are taught to avoid scheduling contract awards in the fourth quarter of the fiscal year, the data in Figure 5 shows that a number of contract awards are, in fact, taking place in the fourth quarter. If unobligated funds can be carried over, because obligation benchmarks have been relaxed, perhaps the PMs are attempting to commit those funds before the end of the fiscal year when the funds’ period of obligation availability ends. Unfortunately, the data provided in the DAMIR data used in Figure 5 does not reveal or imply any causal relationship between the fourth quarter contract awards and the policy memo on obligation rate tenets. More research into the reason for fourth quarter contract awards is needed.

Recent Legislation on Unobligated Balances and Benchmarks

The annual Defense Appropriations Act consistently includes this language: “No more than 20 percent of the appropriations in this Act which are limited for obligation during the current fiscal year shall be obligated during the last 2 months of the fiscal year” (Department of Defense Appropriations Act, 2015). This language is supposed to preclude a rush to award contracts, just to lower unobligated balances at the end of the fiscal year.

Over the past several years, the Congress has marked down the DoD portion of the president’s budget request when large unobligated balances exist. For example, in the DoD Appropriations Act for 2015, the Congress cut $1.76 billion out of the request due to unobligated balances left over from earlier budget years (Congressional Research Service, 2015). Enacting the FY 2012 Defense Appropriations Act, the Congress cut $2.66 billion from the president’s budget request with the rationale that that unobligated balances could be used in lieu of new budget authority. (Congressional Research Service, 2012). Of course,
the problem is that new budget authority has longer periods of availability than old unobligated funds. What might appear to be a rational reduction in new budget authority is really placing a more onerous task on financial managers and contracting officers who must now work more quickly to get the old funds obligated on contracts.

Finally, in the case of recent sequestration cuts, there is a somewhat ironic advantage to having large unobligated balances. For example, in FY 2013, sequestration cuts to DoD investment accounts resulted in an average 11.2% reduction across old unobligated funds. However, the more current FY 2013 budget authority, which has a longer period of obligation availability than old funds, was somewhat preserved, suffering an average sequestration cut of only 5.2% (Congressional Research Service, 2014).

Conclusion

Given the data-driven research just provided, the Under Secretary of Defense (Comptroller) and the Under Secretary of Defense for Acquisition, Technology, and Logistics should exercise caution in implementing their joint policy memorandum that outlined obligation rate tenets. First, recent trends in obligation rates have shown that by the end of the first year of obligation availability, obligation rates for both the RDT&E and the procurement appropriations have dropped below historical norms. Second, this drop in obligation rates has resulted in the carryover of a greater percentage of funds each budget year. Motivated by austere times, the Congress continues to cut the DoD portion of the president’s budget request by amounts nearly equal to the carryover amounts.

Contract awards in the fourth quarter of the fiscal year may indicate that there is still a rush to place funds on contract before the end of obligation availability. While some of these rushed contract awards may represent a good business deal, some of the fourth quarter awards probably waste taxpayer dollars. However, more research is needed into the reasons for fourth quarter contract awards. The number of reprogramming actions and associated amounts of repurposed budget authority may also be an indicator of behavior toward unobligated balances. However, reason codes are missing from recent BTR data, so it was not possible in this research effort to link any change in BTR actions or BTR amounts to the policy memo on obligation rate tenets.

Perhaps the Services/Components are following the joint policy memorandum on obligation rate tenets even though the policy has not been incorporated into any DoD regulation. This may be a causal reason for lower obligation rates and increasing carryover amounts in the RDT&E appropriation accounts. But, why did obligation rates dip five years earlier in the procurement appropriation accounts? Once again, more research is needed.

Whether or not there is any linkage between the data revealed by this research and the joint policy on obligation rate tenets, the Congressional Defense Committees need to be informed of the policy and asked not to decrement RDT&E and procurement accounts without first assessing why large unobligated balances exist for some acquisition programs. Of course the DoD will have to provide more detail on carryover amounts, perhaps as part of budget justification materials, to convince the Congressional Defense Committees that the carryover of funds is legitimate and necessary for program success.

It would be prudent for the Under Secretary of Defense (Comptroller) and the Under Secretary of Defense for Acquisition, Technology, and Logistics to get the Congress to agree with the policy memo on obligation rate tenets before incorporating those tenets into DoD regulations. And, it appears that such a dialog has now begun. In response to a request from the U.S. Senate for ideas on the subject of Defense Acquisition Reform, the
Honorable Frank Kendall III, Under Secretary of Defense for Acquisition, Technology, and Logistics, wrote

For 4 years I have worked to train and encourage our acquisition workforce to take time to get good business deals for the Taxpayer by conducting appropriate upfront analysis, and by doing the systems engineering and planning necessary for successful programs. At the same time our program managers live in a world in which they are punished for not obligating the funds they control on set schedules. We should have realistic plans to execute our budgets, but when a manager has sound reasons to delay obligation, that behavior should not be punished. I have worked with the Under Secretary of Defense (Comptroller) to provide a more balanced approach to how we handle obligation reviews within the Department, and we would like to work closely with Congress in striking a similar balance on this matter. (Permanent Subcommittee on Investigations, 2014, p. 115)

References


Appendix

![Memorandum](image-url)

MEMORANDUM FOR: SEE DISTRIBUTION

SUBJECT: Department of Defense Management of Unobligated Funds; Obligation Rate Tenets

The purpose of this memorandum is to address a long-standing Department of Defense (DoD) problem regarding the way we manage unobligated funds.

The acquisition community, starting with the Under Secretary for Acquisition, Technology and Logistics, has been stressing the importance of spending money in a way that maximizes the value the Department and the taxpayer receive. Our policy encourages managers to obligate funds when a satisfactory contract is negotiated or when they can be used most efficiently. The financial management community, starting with the Under Secretary of Defense (Comptroller), has primarily been measuring program execution against established obligation benchmarks as the basis for sourcing funds for higher Department priorities. Obviously, both goals—the effective acquisition practices and use of resources for the highest Department priorities—have merit. We must strive to meet both goals while also taking into account two types of risks.

First, there are risks in focusing on obligation benchmarks. The threat that funding will be taken away or that future budgets can be reduced unless funds are obligated on schedule is a strong and perverse motivator. We risk creating incentives to enter into quick but poor business deals or to expend funds primarily to avoid reductions in future budget years. We need to rethink how we approach managing mid-year and end-of-year obligations and to change the types of behavior we reward or punish. We will continue to hold our Program/contracts teams accountable for executing to their planned schedules, but we have to stop measuring only benchmark execution as the dispositive method of determining whether funds are available for higher priorities. Such benchmarks should only be a place to start a discussion of obligation management, not a place to end that discussion.

But there are also risks in ignoring obligation benchmarks. For the past several years, Congress has used unobligated balances as a means to reduce our budgets. To avoid this result, we need to stop thinking of the transfer of funds to higher priorities as something we must avoid at any cost. We need to build a culture where maximizing the Department’s buying power for both the taxpayer and the Warfighter as well as meeting the Department’s highest priorities become the primary driving force behind obligation decisions. Often, these funds can be better employed elsewhere and individual programs should not fight to avoid “losing” the unobligated funds.

We believe that the following tenets should be adopted and enforced at all levels of the chain of command, and by acquisition and financial managers throughout the Department:
1. Taxpayer funds should be obligated and ultimately expended only in the taxpayers' interest and if best value is received for the money in support of the Warfighter.

2. While they can be useful indicators, obligation rates slower than established benchmarks should not be the determinative measuring stick for program execution and must not be regarded as a failure.

3. Late obligation of funds should not be presumed to imply that the funds are not needed or that future budgets should be reduced unless there is other evidence to support that conclusion.

4. Providing savings to the organization, military service, or DoD component as early in the fiscal year as possible should be encouraged and rewarded, professionally and visibly.

5. Savings will not be reallocated at any higher DoD level than necessary to fulfill shortfalls in priority requirements.

6. Managers who release unobligated funds to higher priorities will not automatically be penalized in their next year's budget with a lower allocation and may be candidates for additional funding to offset prior year reductions.

This year, the undersigned will begin implementation of a deliberate process by reviewing MDAAPs and other programs that have not obligated funds consistent with normal obligation rates. We will do this together, co-led by officials designated by the undersigned, and we will follow the tenets we have listed. We will start with guidelines for the review, which will be issued soon.

We urge all financial and acquisition managers and the chain of command in each Component to follow the same guiding principles and to implement similar reviews within their organizations.

Robert F. Hale
Under Secretary of Defense
(Comptroller)

Frank Kendall
Under Secretary of Defense for
Acquisition, Technology and Logistics
### Panel 17. Applying Portfolio Techniques to Acquisition Management

**Thursday, May 14, 2015**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:45 p.m. –</td>
<td><strong>Chair:</strong> Joseph L. Yakovac, Jr., LTG, U.S. Army (Ret.), NPS; former Military Deputy to the Assistant Secretary of the Army (Acquisition, Logistics, &amp; Technology)</td>
</tr>
</tbody>
</table>
|               | Peter Vascik, MIT  
|               | Adam M. Ross, MIT  
|               | Donna H. Rhodes, MIT |
|               | **An Analytic Model for DoD Divestments**  
|               | Lisa Oakley-Bogdewic, MITRE Corporation  
|               | Emaan M. Osman, MITRE Corporation  
|               | **Portfolio Acquisition—How the DoD Can Leverage the Commercial Product Line Model**  
|               | Peter Modigliani, MITRE Corporation |

**Joseph L. Yakovac, Jr., LTG, U.S. Army (Ret.)**—retired from the United States Army in 2007, concluding 30 years of military service. His last assignment was as director of the Army Acquisition Corps and military deputy to the assistant secretary of defense for acquisition, logistics, and technology. In those roles, Lt. Gen. Yakovac managed a dedicated team of military and civilian acquisition experts to make sure America’s soldiers received state-of-the-art critical systems and support across a full spectrum of Army operations. He also provided critical military insight to the Department of Defense senior civilian leadership on acquisition management, technological infrastructure development, and systems management.

Previously, Lt. Gen. Yakovac worked in systems acquisition, U.S. Army Tank-Automotive Command (TACOM), and in systems management and horizontal technology integration for the Office of the Assistant Secretary of the Army for Acquisition, Logistics, and Technology. He has also served as executive officer and branch chief for the Bradley Fighting Vehicle and as a brigade operations officer and battalion executive officer, U.S. Army Europe and TACOM.

Lt. Gen. Yakovac was commissioned in the infantry upon his graduation from the U.S. Military Academy at West Point. He served as a platoon leader, executive officer, and company commander in mechanized infantry units. He earned a Master of Science in mechanical engineering from the University of Colorado at Boulder before returning to West Point as an assistant professor.

Lt. Gen. Yakovac is a graduate of the Armor Officer Advanced Course, the Army Command and General Staff College, the Defense Systems Management College, and the Industrial College of the Armed Forces. He has earned the Expert Infantry Badge, the Ranger Tab, the Parachutist Badge, and for his service has received the Distinguished Service Medal, the Legion of Merit three times, and the Army Meritorious Service Medal seven times.
Abstract

The modern warfighter operates in an environment that has dramatically evolved in sophistication and interconnectedness over the past half century. With each passing year, the infusion of ever more complex technologies and integrated systems places increasing burdens on acquisition officers to make decisions regarding potential programs with respect to the joint capability portfolio. Furthermore, significant cost overruns in recent acquisition programs reveal that, despite efforts since 2010 to ensure the affordability of systems, additional work is needed to develop enhanced approaches and methods. This paper discusses research that builds on prior work that explored system design tradespaces for affordability under uncertainty, extending it to the program and portfolio level. Time-varying exogenous factors, such as resource availability, stakeholder needs, or production delays, may influence the potential for value contribution by constituent systems over the lifecycle of a portfolio, and make an initially attractive design less attractive over time. This paper introduces a method to conduct portfolio design for affordability by augmenting Epoch-Era Analysis with aspects of Modern Portfolio Theory. The method is demonstrated through the design of a carrier strike group portfolio involving the integration of multiple legacy systems with the acquisition of new vessels.

Introduction

Enabled by the emerging widespread availability of high speed computing, computational Tradespace Exploration (TSE) has become a valuable tool. TSE empowers system engineers to consider far more potential designs than could be done through prior Analysis of Alternatives methods (Ross & Hastings, 2005). A recently developed approach,
Epoch-Era Analysis (EEA), enables the conceptual design of systems that are resilient to potential changes in context and needs (exogenous uncertainties) over the lifecycle of the system. EEA enables quantitative support for the design of particular time-contingent system properties such as survivability, flexibility, and affordability (Ross et al., 2010).

Since the issuance of the 2010 memo Better Buying Power (BBP) and the implementation of the Department of Defense (DoD) Efficiency Initiative (Carter, 2010), a significant body of research has been generated to support design for affordability, including advanced TSE techniques. To this end, EEA was adapted for affordability analysis in naval acquisitions by Schaffner et al. (2013) through the introduction of aggregate cost and schedule considerations. As system interconnectedness and interdependence continues to increase, especially in Navy and DoD operations involving numerous assets, conceptual design techniques that consider only the acquisition and operation of individual systems are not fully sufficient. Of additional value in this regard is expanding the scope of such techniques to support acquisition decisions at the multi-system level; this necessitates consideration of two related concepts: systems of systems (SoS) and portfolios.

A SoS is a dynamic network of constituent systems that exhibit varying levels of operational and managerial independence, but operationally interact so as to achieve mutually desired, oftentimes emergent, capabilities (Maier, 1999). A portfolio is a construct which describes a collection of assets, acquisition programs, and research programs that are jointly invested in to exploit qualities of the set, regardless of whether the assets are operationalized independently or participate in a SoS. The Carrier Strike Group (CSG) case study conveniently illustrates the difference between SoS and portfolio design. The CSG portfolio design problem seeks to ascertain what acquisition strategies result in an affordable set of available assets that may be assembled into any number of CSG SoS to meet a variety of possible desired performance attributes. The SoS design problem, on the other hand, would occur downstream of portfolio design where a designer seeks to select available assets from the portfolio, apply concept of operations (CONOPS) to dictate SoS interactions, and meet specific desired capabilities.

SoS and portfolio design considerations represent unique challenges for tradespace analysis, and especially for affordability analysis under uncertainty. Because engineering portfolios may include assets that are SoS themselves, and because there is typically a high degree of interaction and interdependencies in the costs, risks, and capabilities of the assets, traditional portfolio assessment techniques must be modified to address these complexities that violate prior assumptions. SoS-based design approaches enable consideration of these unique qualities of engineering portfolios and shall inform the development of the method presented in this research.

This paper discusses recent efforts to adapt a resource-centric approach to EEA developed by Schaffner, Ross, and Rhodes (2014) for use in affordable portfolio design through the integration of elements of Modern Portfolio Theory (MPT) and the SoS design literature. The proposed method is demonstrated in a case study on the acquisition of a portfolio of assets from which a CSG may be assembled. The method is leveraged to provide design insight into CSG lifecycle affordability. This insight is achieved by identifying the utility and costs of portfolio designs in potential future contexts that embody a variety of uncertainty factors, such as unit availability, budget constraints, capability requirements, strategic threats, and technology development. The 11 processes of the proposed method are described as applied to the CSG portfolio, and they illustrate the method’s potential value for naval acquisition and operations. Furthermore, a discussion of the potential for broad applicability of the method to other DoD capability portfolios is proffered, specifically with respect to the Joint Capabilities Integration and Development System (JCIDS).
**Motivation**

Between 1997 and 2011, there were 74 Nunn-McCurdy program unit cost breeches in 47 of 134 major DoD acquisition efforts (Government Accountability Office [GAO], 2011). According to an audit by the GAO, many of these breeches corresponded to context changes in the environment surrounding the acquisition programs. These context changes included affordability measurement statute modifications, presidential administrations turnover, unit order size reduction, schedule changes, and requirements changes. These cost breeches occurred in over one third of major DoD acquisition programs and indicate the need for conceptual design methodologies that consider potential changes in context in order to achieve consistent lifecycle program affordability.

In light of these breeches, and supported by the DoD emphasis of an “affordability mandate” in BBP 1.0, 2.0, and now 3.0 (Carter, 2010; Kendall, 2012, 2014), design for affordability literature and practice grew rapidly. This provides the foundation for extending EEA to the portfolio-level. First, the concept and meaning of the term *affordability* was explored and defined by many organizations including INCOSE and the NDIA, as well as by the DoD (Schaffner et al., 2013). Second, metrics were developed to assess the contextual and dynamic attributes of affordable systems (Bobinis et al., 2013), and a variety of tools were produced to support affordability tradeoffs between potential systems. Third, EEA was employed to reveal system affordability across a variety of uncertain futures and to provide insight into changing contexts, such as those which caused many of the recent Nunn-McCurdy breeches (Wu, Ross, & Rhodes, 2014).

These advances, however, address only a part of the challenge in achieving affordability for the DoD since they do not explicitly consider the higher order complexities inherent to multi-system acquisition and operations (Wu, 2014). As early as 2010, DoD decision makers recognized that design for affordability at the system level did not necessarily translate to the affordability of the overall capability portfolio. Remarks by General Peter Chiarelli indicate an understanding of this concept and call for a portfolio-level conceptual design paradigm (Association of the United States Army, 2010, p. 1):

> If you look at any one of these systems as an individual system, you can sell just about anything. But, when you look at [an] entire portfolio you can start to see where we have duplication in different systems or maybe we’re overinvesting in one and underinvesting in another.

A variety of portfolio management techniques have been developed to begin to fill this design and acquisition capability gap. The *Systems Engineering Guide for System-of-Systems* (2008) provides guidelines for the engineering of SoS in DoD acquisitions. Komoroski et al. (2006) applied a variant of real options analysis to identify long-term SoS acquisition strategies for information technology. The Computational Exploratory Model by Mane and DeLaurentis (2011) was developed to assess development networks of SoS architectures. Initial efforts to inform the acquisition and integration of systems in a SoS through MPT were also made by Davendralingam, Mane, and DeLaurentis (2012). Epoch-Era Analysis complements these previous efforts by adding the ability to assess the influence of changing contexts on the affordability of potential portfolios.

**Extending EEA for Affordability From “System-Level” to “Portfolio-Level”**

To clarify the terminology used in this paper and link it to trends in the SoS literature, a set of terms is introduced to describe three “tiers” of design abstraction. Table 1 presents an example of the design scope for each tier of design abstraction as applied to the CSG case.
System-Level: System design occurs at a “level of decomposition that is inclusive of a major architectural element and is semi-independent from the rest of the architecture” (Walton, 2002). The designer of a system typically has full design authority, and although a system may be composed of multiple components (such as a launch vehicle), it is not considered a SoS as the sub-elements are not managerially independent (Maier, 1999).

Program-Level: Program design is distinguished from system design in that it requires joint consideration of multiple independent or semi-independent constituent elements (typically systems themselves). However, like system-level analysis, the designer of a program is typically assumed to have a moderate to high degree of design authority. Two primary types of programs have been identified, and they are distinguished by the attributes of the constituent systems:

- **Type I programs** are composed of homogenous constituent systems. Type I program design is readily conducted through traditional EEA where the most promising single system solution is also expected to produce the greatest program benefit. Initial research into Type I programs was conducted by Wu (2014).
- **Type II programs** are composed of heterogeneous constituent systems which often complete similar missions and are evaluated by a common set of value metrics. Because Type II programs concern either semi-independent or fully-independent constituent systems, as opposed to the closely managed homogenous systems in Type I programs, the design problem involves SoS challenges.

Portfolio-Level: A portfolio is a collection of selected assets that may be either new or legacy programs as defined above, which are simultaneously invested in to collectively provide a set of capabilities. A portfolio designer does not necessarily have a significant level of control over the design of the constituent programs, or their ultimate operationalization in a SoS, but can create a portfolio with attractive procurement, management, and capability features based upon the possible assets and their likely applications. Design at the portfolio-level must not only consider traditional financial portfolio investment techniques that identify emergent properties from a set of independent assets, but also must consider SoS techniques that consider value arising from the potential interaction of the assets when operationalized.

Table 1. Design Abstraction Tiers as Applied to a Carrier Strike Group

<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>PROGRAM</th>
<th>PORTFOLIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Unit</td>
<td>Anleigh Burke-class destroyer</td>
<td>Carrier strike group</td>
</tr>
<tr>
<td>Design Goal</td>
<td>Meet required capability criteria for assigned mission</td>
<td>Provide multi-mission offensive and defensive capabilities</td>
</tr>
</tbody>
</table>

In prior research, Wu (2014) proposed a methodology for program and portfolio-level affordability analysis utilizing a *bottom-up “survival of the fittest”* approach. Wu’s approach leveraged EEA to identify promising individual systems for affordability; such system-level analysis is an extensively explored capability of EEA. At the next tier of design abstraction,
the program level, new program-level performance attributes were applied to a set of programs seeded from the “promising” systems identified in system-level design. It was then ascertained which combination of these systems most promisingly fulfilled the program-level goals. This “wrapping” approach was a logical first step towards program and portfolio design with EEA. However, the method may prematurely discard potentially valuable designs that do not appear attractive in lower levels of design abstraction, but produce a favorable overall portfolio.

The concern for premature abandonment of potential assets is supported by Walton (2002). Walton’s work applied portfolio theory to tradespace analysis for a space science mission and found that the most promising portfolio of assets to minimize uncertainty was not consistently a constellation of the most promising single satellite design(s). Rather, Walton found that designs for portfolio-level minimum uncertainty included sub-par, system-level solutions that interacted to exhibit emergent benefits and the highest portfolio-level utility.

Recognizing the limitations of both the bottom-up and top-down (traditional requirements-driven design) approaches to portfolio design for affordability, Wu suggested that a two pronged method to leverage the strengths of each approach may lead to higher-utility portfolios. TSE is inherently a bottom-up process where constituent systems are first designed for desired system level attributes, and then linked together through multiple levels of larger subsystems to result in portfolio capabilities. MPT is an inherently top-down approach where designers identify desired portfolio-level attributes (big picture), and an algorithm then seeks to compile a set of investments (systems) which satisfy those attributes in aggregate (Amenc & Sourd, 2005). This paper intends to leverage both of these techniques by applying elements of MPT and TSE to EEA. Synthesized bottom-up and top-down models have found traction in land use portfolio planning and energy portfolio management (Castella et al., 2007; Wing, 2006).

Overview of an EEA-Based Approach for Design for Affordability

Acquisition program planning under affordability considerations requires analysis of factors such as system development schedules, legacy hardware and operations, resources, and political capital. Furthermore, these factors must be considered with respect to the dynamic environments in which they exist over the system lifecycle. Epoch-Era Analysis is an effective mechanism to evaluate acquisition strategies in anticipation of future context shifts. An epoch is a time period of static context and stakeholder expectations, like a snapshot of a potential future. An era is an ordered sequence of epochs with finite durations that describe a potential progression of contexts over the lifecycle of the system.

EEA consists of several distinct, but related analysis techniques. In single-epoch analysis, potential portfolios are evaluated in individual epochs to determine how close to the Pareto front the portfolio lies. When the performance of the same portfolio is compared in different epochs through multi-epoch analysis, the various influences of contextual uncertainties may be perceived and shifts in portfolio proximity to the Pareto front may be observed. If ordered epochs are strung together into an era, changes in the value of proposed portfolios over time becomes apparent, and portfolios which may become unaffordable are identifiable. Figure 1 illustrates how single-era analysis may reveal lifecycle deficiencies in initially promising portfolio designs.
Note. This Epoch-Era Analysis reveals the performance (including affordability) of a system through a sequence of varying contexts illustrating the potential lifecycle value of the system.

**Figure 1. Epoch-Era Analysis**  
(Ross & Rhodes, 2008)

In prior research, Schaffner et al. (2014) developed a composite method for affordability based upon the Responsive Systems Comparison (RSC) method, and applied their approach to the acquisition of a Next Generation Combat Ship (NGCS) for the Navy. RSC is a variant of EEA which leverages TSE as a powerful tool to conduct analysis. Complex system design, and particularly SoS design, involves multiple dimensions of relevant benefits, expenses, and boundary conditions which do not lend themselves to optimization and are often too complex for intuitive decision making. TSE applies the capabilities of modern computing to enumerate a large variety of design alternatives to support a decision maker to holistically investigate subtle tradeoffs and previously unconsidered designs. As a result, RSC may lead a designer to select an acquisition program with superior lifecycle results than those determined by a numerical optimizer operating under simple tradeoffs and rules of thumb (Wu et al., 2014). Building on this prior research, this paper describes the extension of the method to the program and portfolio levels.

**Overview of Modern Portfolio Theory in Systems Engineering**

Modern Portfolio Theory (MPT) has a ubiquitous impact on financial investment strategies. MPT provides a methodology to identify an efficient frontier of portfolio investments based upon the elicited values and preferences of the investor(s) and the asset performance forecasts. Such an efficient frontier is composed of potential portfolios of investments which maximize the return on investment while minimizing the risk. To achieve this end, MPT relies on the concept that groups of investments with negative trending covariance exhibit a portfolio risk which is less than the average of the risks of each constituent investment (Amenc & Sourd, 2005).

Numerous derivatives have extended the capabilities of MPT. Post-Modern Portfolio Theory (PMPT) allows for the consideration of non-normally distributed risks (a more realistic assumption) and provides for the minimization of “downside risk” (negative outcomes) rather than mean variance (Swisher & Kasten, 2005). Walton (2002) employed TSE to allow for formal tradeoffs between value and uncertainty (risk) in an effort to reveal
“synergistic combinations of architectures.” Furthermore, Walton’s work introduced semi-variance as a method to treat upside and downside risks independently in portfolio optimization. The authors have elected to utilize MPT in this initial application to EEA due to its greater simplicity, and the availability of literature documenting application to fields beyond finance.

The fusion of aspects of portfolio theory with TSE is not a novel concept. Davendralingam et al. (2012, p. 63) introduced a modification to a model of SoS acquisition originally developed by Lane et al. (2010) as a mechanism to support portfolio design and “maintain compliance with the ‘top-down integration, bottoms-up implementation’ paradigm” of traditional SoS design. Davendralingam et al.’s approach relied upon an ongoing, iterative circuit of design, implementation, and feedback to mitigate dynamic contextual uncertainties. The method in this paper utilizes EEA to consider the potential uncertainties of the internal and external environment at the outset, and select a portfolio which is resilient against changes in context.

**Generalization of MPT and Combination with EEA**

While MPT and EEA share a variety of commonalities including value elicitation from stakeholders/investors, the use of models to describe investment/system value, and a foundation in utility theory, there are also fundamental differences in the design of SoS which violate assumptions necessary for MPT. Ricci and Ross (2012) conducted a review of the similarities and differences of MPT and EEA. Their work yielded Table 2, which describes key differences between financial and SoS portfolios.

### Table 2. List of Salient Differences Between Financial and SoS Portfolios

(Ricci & Ross, 2012)

<table>
<thead>
<tr>
<th>Financial Portfolios</th>
<th>SoS Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Often assume normally distributed asset returns</td>
<td>Performance models use various distributions</td>
</tr>
<tr>
<td>Linear aggregation of asset return</td>
<td>Non-linear relationships dictate asset aggregate performance</td>
</tr>
<tr>
<td>Covariance can express correlation of assets</td>
<td>Covariance is not sufficient to describe correlation</td>
</tr>
<tr>
<td>Various types of assets are often generally available</td>
<td>Assets may be limited based on economics, technology, politics</td>
</tr>
<tr>
<td>Availability of assets remains fairly constant</td>
<td>Assets may change availability over time</td>
</tr>
<tr>
<td>No portfolio diversification cost</td>
<td>Diversification costs</td>
</tr>
<tr>
<td>Low carrying costs</td>
<td>Possibly large carrying costs</td>
</tr>
<tr>
<td>Small switching costs</td>
<td>Possibly large switching costs</td>
</tr>
</tbody>
</table>

Ricci and Ross propose a variety of modifications to MPT which address asset availability, diversification costs, carrying costs, and switching costs. However, the disconnect between SoS and financial portfolios over non-normal distributions and non-linear relationships of constituent systems to portfolio value is yet to be addressed.

In MPT, the distribution of return (utility) is assumed to be a normal distribution around the expected value. This is an effort of MPT to characterize asset response to potential changes in future context. The assumption of a normal distribution may not be appropriate for engineering portfolios where, especially for novel systems, there is not a large set of historical data upon which constituent system performance simulations may be grounded. EEA proves a convenient mechanism to address this challenge. Multi-Epoch and era-level analyses reveal changes in utility through possible changes in context and expectations, while also identifying path-dependent uncertainty. EEA may therefore be used to determine promising portfolios that maintain acceptable value across these potential futures, rather than relying on value predictions from the aggregation of utility distributions. In a sense, when EEA and MPT are combined, assumptions about the distribution of risk of...
constituent systems no longer need to be made, but rather the impact of various uncertainty factors may be readily simulated and examined in the ultimate design choices.

When considering an engineering portfolio, the utility (and expense) of the overall portfolio may not simply be the aggregated, linear-sum of the stand-alone utilities of the constituent elements. Rather, engineering portfolio utility may be greater or less depending upon additional operational relationships among the constituent elements. For example, while a Ticonderoga-class cruiser provides a great deal of anti-ship missile defense value, two cruisers do not necessarily provide a potential CSG with twice the value. The anti-missile systems, radar systems, and concept of operations (CONOPS) used by the cruisers are identical (and therefore susceptible to the same vulnerabilities) and would not provide a CSG with twice the value of a single cruiser. Therefore the presence of both systems should not increase the acquisition portfolio value for these capabilities by the linear sum of both systems’ utility. Conversely, the anticipated Air and Missile Defense Radar (AMDR) on Flight III Arleigh-Burke destroyers is expected to provide enhanced value beyond the linear aggregation of the individual ship capacities when two more ships operate in conjunction. While MPT does not consider such investment interactions, the development of SoS utility from the capabilities of constituent systems is an active field of research. Therefore, this paper adopts two possible constructs based on SoS research as initial attempts to address this challenge for portfolio design: the capability tree and combination coefficients.

**Introduction of the Method for Portfolio Affordability Analysis**

Portfolio-Level Epoch-Era Analysis for Affordability (PLEEAA) is proposed as a new method for portfolio affordability analysis. The key innovative feature of the method is the fusion of elements from Modern Portfolio Theory with the capabilities of Epoch-Era Analysis. Figure 2 displays a graphical representation of how this is achieved.

---

**Figure 2. Portfolio-Level Epoch-Era Analysis for Affordability (PLEEAA) Process Model for Affordability Analysis**

Multi-Attribute Tradespace Exploration (MATE) is a method that has been used extensively in system-level design. When used with EEA, MATE may design for a variety of “ilities,” including affordability. As shown in RSC, MATE combined with EEA may be used to consider dynamic uncertainty of contexts and needs. PLEEAA extends RSC with two new constructs to enumerate viable portfolios: a portfolio selector and a portfolio design tool. Additionally, a “portfolio capability tree” is also implemented to link system performance to portfolio utility.
Portfolio Design Tool and Portfolio Selector

A fundamental element of MPT is “asset allocation,” or the identification of potential asset classes which a portfolio designer (the systems engineer applying PLEEEAA) desires to consider for inclusion in the portfolio. In finance, these assets may be stocks, bonds, or cash (Amenc & Sourd, 2005). For Navy portfolios, asset allocation may involve identifying fundamental categories of assets such as cruisers, destroyers, aircraft, and submarines, for example. The portfolio designer may apply specific constraints to each class, such as a maximum number of units (or funds) they are willing to allocate to that class; these class constraints mirror investment thresholds in financial portfolios.

The asset allocation decisions in the portfolio design tool represent the most significant lever a portfolio designer has to influence the outcome of the PLEEEAA method. If a portfolio designer wishes to constrain the analysis to reflect current institutional inertia or design paradigms, they may develop highly specific classes and class constraints which force the portfolio selector to only consider designs similar to existing portfolios. However, if the designer wishes to explore potentially novel approaches and unconsidered emergent qualities of constituent systems, they may choose to set few to no class constraints on the analysis. This would enumerate a far greater portion of the potential portfolios design space, but possibly at significant computational cost.

A second role of the portfolio design tool and class constraints is to reduce the computational complexity of the portfolio design problem, as shown in the case example.

Portfolio Capability Tree

A key challenge of any tradespace-based portfolio analysis is linking constituent system performance attributes with portfolio-level capabilities. This research uses the concept of a capability tree, a capability-based value mapping, to both percolate portfolio designer needs down through the portfolio levels to the constituent system managers, and then to amalgamate system-level performance attributes back up the capability tree and define a portfolio-level utility. The structure and inspiration of capability trees are rooted in means-ends objectives networks, as developed by Keeney (1992). Means-ends objectives networks link decision alternatives to their impact on the overall objective to enable quantitative modeling. Capability trees seek to link portfolio asset options to the overall portfolio utility for the same purpose.

The process to develop the capability tree begins with value-elicitation. A portfolio designer decomposes the strategic objectives of the portfolio-level stakeholders into a set of desired capabilities, or performance attributes. Performance attributes become a measure of how well a potential portfolio meets the needs of the portfolio designer. Each performance attribute of the portfolio is communicated to a unique program-level manager who is the decision maker for constituent programs contributing to that portfolio performance attribute. Similar to before, the program-level manager decomposes their values into a set of program-level performance attributes which are communicated to either another program-level manager, or to the constituent systems of the portfolio.

It is important to recognize that the portfolio designer and program-level managers at each level in the portfolio hierarchy are empowered to develop their own performance attributes and mental value models to measure the operation and utility of the systems (or programs) in their subtree. The utility measurement developed is then shared by the in-level manager with the manager one level above in the capability tree hierarchy. Therefore, the utility measurement of a program must be effectively communicated at the interface of the two managers, referred to as a “node.” The effective communication of utility between individuals is a non-trivial task, and in general, may benefit from multi-party utility
negotiations such as those under research by Fitzgerald and Ross (2015). The capability tree model, as described here, is an imposed constructed value model and is illustrated in Figure 3, where nodes are represented by collate symbols.

The “nesting” process utilized to decompose portfolio-level strategic objectives into performance attributes, and communicate these values to constituent programs or systems, creates a two-dimensional root or tree structure of desired performance attributes at various levels of the portfolio hierarchy. The multi-document symbol in Figure 3 represents a program-level manager who utilizes their own value and performance models to determine the utility of the constituent systems or programs in their subtree.

![Figure 3. PLEAA “Capability Tree” Portfolio Design Architecture](image)

Through the capability tree, desired portfolio strategic objectives flow down to lower level program managers as performance attributes. The capability trees are extendable to multiple levels of programs: Three have been shown in Figure 3. Each additional level of the portfolio hierarchy allows for the design of more intricate portfolios through TSE techniques, but may also substantially increase the analysis complexity. All “branches” of the tree do not need to have the same number of levels. Some capabilities will naturally terminate in system-level attributes after the first level, while others may require multiple levels of decomposition to equate to system-level performance attributes. A portfolio objective is fully
decomposed when the system-level “leaf node” performance attribute is readily described by a metric of the potential component systems, such as the number of missiles on a ship.

In the command hierarchy of military institutions, utility handoffs may be straightforward due to the subordinate decision-making (chain of command) architecture. For example, the structure of the Joint Capabilities Integration Development System (JCIDS) represents the information hierarchy that the capability tree seeks to leverage for portfolio affordability analysis. JCIDS manages the DoD’s joint capability portfolio and conducts capability gap assessments, among other capability requirement development and approval functions, by identifying joint capability areas (similar to the capability tree portfolio-level performance capabilities) and engaging decision makers in a four-tier hierarchy which moves information from individual subject matter experts up to four star decision makers (Chairman of the Joint Chiefs of Staff, 2015b). In a very real sense, the PLEEAA method adopts the portfolio management structure of the JCIDS process established by Chairman of the Joint Chiefs of Staff (2015a) and integrates MPT and EEA tradespace exploration techniques to produce more robust, data-driven portfolio analyses.

**Complementary and Substitute Systems**

A fundamental challenge unique to engineering portfolios versus financial portfolios is that portfolio-level capability may not simply be an aggregate sum of the constituent asset capabilities. As discussed previously, while a portfolio is a non-operational construct that describes the acquisition and inherent values of a set of assets, engineering portfolios must consider emergent value from asset interaction in a SoS. The concept of complementary and substitute systems in SoS has been an area of intense academic focus. To provide a few definitions,

1. Complementary systems are two or more constituent systems that experience a change in value delivery towards existing performance attributes, or gain capability in new performance attributes, when simultaneously present in a SoS (in this case, a portfolio). Such changes in value usually result from an adjustment in the CONOPS of the systems. The sign and magnitude of the performance change is variable. For example, the missile strike capability of a submarine may be significantly increased when combined with the advanced radar and fire control capabilities of a missile cruiser. However, the same submarine may experience a reduction in stealth capability as it must be within a certain range of the cruiser and transmit targeting information.

2. Substitute systems are those that provide an overlapping performance attribute capability in an operating scenario (i.e., CONOPS) of interest. A guided missile cruiser and guided missile destroyer may perhaps be considered substitute systems in terms of anti-ballistic missile capabilities. However, they would not likely be substitutes for littoral operations as their capabilities and vulnerabilities differ.

The PLEEAA capability tree architecture provides a unique mechanism to identify and assess substitute systems. Since all potential systems are simultaneously evaluated by each bottom-level portfolio hierarchy manager, substitute systems that provide similar capability are likely to be identified. The manager’s value model may then appropriately determine the node’s aggregate utility resulting from the interacting substitute systems.

The identification of complementary systems is not as straightforward, but is still enabled by the capability tree architecture. Again, because bottom-level portfolio hierarchy managers ideally have visibility of the relevant potential constituent systems of the portfolio.
that contribute to the performance attribute of interest, they will likely be able to identify system pairings which complement each other’s value delivery. The bottom-level managers, unlike the portfolio designer or higher level program managers, will reasonably have the expertise to understand the potential interactions between systems. In other words, the capability tree framework releases portfolio-level designers from trying to make technical evaluations on the numerous constituent systems in the portfolio, but rather allows each manager in the capability tree to only consider system or program interactions and values at their node. If a bottom-level portfolio manager is unable to identify system interactions, then emergent complementary value may not be properly considered in the portfolio design.

On an additional note, both complementary and substitute systems are likely to represent opportunities for cost savings through joint development or production programs. While this information is unlikely to be offered by the constituent system operators during value elicitation, the bottom-level managers will likely recognize such potential. Therefore, the capability tree framework may allow for the adjustment of both utility and costs with respect to complementary and substitute system interactions.

While the capability tree framework may help discern the influence of complementary and substitute systems, the approach outlined above would require the decision-making models of each manager to be well represented at every node. For this initial CSG case study, and indeed for many conceptual design or acquisition programs, this assumption is not realistic. Therefore, this research also adopts a second approach to manage complementary and substitute systems in lower fidelity analyses.

In her work to apply tradespace exploration to SoS, Chattopadhyay (2009) introduced the concept of “level of attribute combination complexity.” Chattopadhyay identified three levels of combination complexity which express general, first-order estimates of performance attribute interaction from different constituent systems in the portfolio. “Low-level combination” characterizes performance attributes which exist independent of one another. An example of low-level combination may be an E-2 Hawkeye early warning capability and a Littoral Combat Ship (LCS) green-water minesweeping capability. “Medium-level combination” describes attributes of systems which contribute to the delivery of the same portfolio performance attribute, but are characterized by a sharing of value deliver, such as through a handoff of responsibility. The handoff of targeting information from a Ticonderoga cruiser to a Virginia-class submarine may be envisioned as medium-level combination for the cruise missile strike performance attribute. Finally, “high-level combination” are system attributes which interact to simultaneously provide a portfolio performance attribute, such as the anti-submarine warfare capabilities of a Virginia-class submarine and an Arleigh-Burke destroyer. A given system may contribute to performance attributes at one or all of these levels of combination complexity when considering their interactions with other systems.

PLEEAA elicits bottom-level portfolio hierarchy managers to characterize the combination complexity for each performance attribute in their area of responsibility for each of the potential constituent systems. Various models may be used to represent the impact of the level of combination on the resulting utility and cost of the portfolio. For the CSG case study, in general, low-level combination attributes are not adjusted, medium-level combination attributes utilize alternative models for utility calculation, and high-level combination attributes receive appropriate utility and cost multipliers.

**Overview of PLEEAA**

PLEEAA is a variant of the RSC method for affordability analysis developed by Schaffner et al. (2013). The Gather–Evaluate–Analyze structure of the RSC method was
maintained and supplemented with additional steps necessary for portfolio design. Specifically, the value-driven design formulation now occurs with multiple groups of stakeholders corresponding to different levels of the portfolio hierarchy.

The basic steps of PLEEAA are represented as an 11-process method shown in Figure 4. Processes 1 through 5 involve value elicitation of multiple-levels of stakeholders in the portfolio to define the problem scope, assess stakeholder needs, identify contextual variables, and assess combination complexity information for potential systems. Process 6 conducts the composite EEA/MPT analysis to produce a tradespace of potential portfolio designs for the considered epoch. Before proceeding to the following processes, feedback is provided to the designers and stakeholders to allow for adjustments in the provided information, as necessary. This feedback loop is a key element of TSE as stakeholder values may change as portfolio options and tradeoffs are made clear. Finally, Processes 7 through 11 support designers at the portfolio-level to compare the dynamic properties of potential portfolios in light of their anticipated performance in a variety of point futures (epochs), as well as possible lifecycle narratives (eras). These processes are described and applied to the CSG study in the following section.

**Figure 4. A Graphical Overview of the Modified Gather–Evaluate–Analyze Structure for PLEEA**

**Demonstration Case: Carrier Strike Group Design for Affordability**

According to a 2006 RAND study, the cost growth rates for new naval units such as “amphibious ships, surface combatants, attack submarines, and nuclear aircraft carriers have ranged from 7 to 11 percent,” an inflationary rate which significantly outstrips development costs in other sectors (Arena et al., 2006). In the decade since this report, the severe cost overruns in the Littoral Combat Ship (LCS) and Zumwalt-class destroyer (DDG1000) programs have likely exacerbated this figure. In light of this matter, it appears...
appropriate to apply PLEEEAA to an application of large ship acquisitions in order to display the potential of this method to assist decision problems in the following ways:

- As the LCS and DDG1000 programs are scaled back (or eliminated) and replaced with alternative systems, portfolio-level affordability analysis may aid the identification of resulting capability gaps in the Navy’s strategic portfolio.
- The use of a PLEEEAA may facilitate the definition of acquisition program requirements which limit capability creep, such as what was seen in the LCS development.
- Portfolio-level affordability analysis will give decision makers insight into the value tradeoff of investing in high-capability, high-costs systems, versus low-capability, low-cost systems for future naval group operations.
- EEA, when applied to a ship acquisition portfolio, will allow designers to foresee the affordability of the proposed capability portfolio in multiple potential future scenarios.

Considering these potential benefits to the naval acquisition process, the design of a portfolio of systems, from which a CSG(s) may be developed, presents itself as an appropriate case study to demonstrate PLEEEAA. Beyond the relevance of the analysis to the DoD’s goals of an affordable Navy and the emergence of new tactics and asymmetric threats for exiting CSGs, the complexity the CSG design problem is conveniently reduced through a series of realistic assumptions. The simplifying assumptions adopted include the following:

1. A CSG is a directed SoS as defined by Maier (1999), in which a central authority (the combatant commander [CCDR] and operational commander) have decision authority over the constitute systems (Chief of Naval Operations, 2010). A directed SoS is ideal for the initial application of this method as the additional complexities inherent to incomplete managerial influence or decentralized control are avoided (Shah, 2013).

2. The baseline composition and basic mission capabilities of CSGs are well-defined by Chief of Naval Operations (2010), enabling effective bounding of the potential portfolio asset set for the purposes of this case study. Furthermore, a CSG is intended to function autonomously for many of its operations. This further simplifies the scoping of the portfolio boundary conditions and interfaces.

3. The hierarchy of the Navy designates specific decision-making authority to specific individuals. This structure, including the Navy’s use of subject matter experts, is directly paralleled by the capability tree structure of PLEEEAA. Therefore, in the analysis, each node of the capability tree may be mapped to a specific decision maker in the CSG command structure.

The following sections briefly describe each process in PLEEEAA. A representational outcome is included, which provides a first pass, high-level application of PLEEEAA to CSG portfolio design, subject to the assumptions highlighted previously in this section. The values of stakeholders, models for performance aggregation, and system performance parameters were notionally created from publically available information and feedback from experts familiar with the systems.

**Process 1: Value-Driving Context Definition**

The first process in PLEEEAA begins by identifying the basic problem statement and design space for the proposed portfolio. The portfolio-level stakeholders are identified and
engaged as necessary to formulate relevant exogenous uncertainties and outline initial value propositions. An inchoate set of potential constituent systems is constructed, and the portfolio designer’s degree of influence over these systems is predicted.

*Representational Outcome:* According to the Chief of Naval Operations (2010), the primary portfolio-level stakeholders influencing the design of a CSG are the combatant commander (CCDR) and operational commander of the naval group. The CSG value proposition is outlined therein as a “responsive, flexible capability for sustained maritime power projection and combat survivability to shape the operation environment, respond to crises, and protect the United States and allied interest in any threat environment.” An initial set of 12 potential constituent systems is provided in the work instruction. Seven more systems were added to capture the NGCS work conducted in prior research and the new Zumwalt-class destroyer. For this analysis, the portfolio designer is assumed to have total control over the acquisition of new systems for the CSG portfolio (subject to the defined constraints).

**Process 2: Portfolio-Level Stakeholder Value-Driven Design Formulation**

In Process 2, the portfolio designer elicits a variety of information from the portfolio-level stakeholders to establish the root (top level) of the capability tree and clarify the constraints of the portfolio design and composition.

- **Performance Attributes:** A set of overarching capabilities which the portfolio must be able to fulfill to meet the strategic objectives. These capabilities, or performance attributes, are assigned weights based on elicited information to reflect the stakeholder preferences. A utility function is developed.

- **Expense Attributes:** Portfolio resource statements are translated into specific costs measured in the portfolio. Acceptable expense thresholds are identified, and expense functions are created.

- **Portfolio Investment Strategy Constraints:** Constraints are set concerning viable portfolio composition. These constraints may define limitations on resources, types of constituent systems, and acceptable risk. They could also be more specific and govern internal investment strategy decisions such as the maximum number of component systems allowed or the minimum resource allocation value to any single system.

*Representational Outcome:* Simplified interpretations of notional, portfolio-level performance and expense attributes for a CSG portfolio are shown in Table 3. For this initial case study, the performance and expense attributes yield multi-attribute utility and multi-attribute expense through linear weighted sum aggregation functions. The value weights at each node sum to one. This requires an assumption that the performance and expense attributes contribute independently to aggregate value. While this is not necessarily a realistic assumption for a CSG portfolio, it is sufficient first order estimation for the demonstration purposes of this case study.
Table 3. Portfolio-Level Performance and Expense Attributes for a CSG

<table>
<thead>
<tr>
<th>Performance Att.</th>
<th>Notional portfolio strategic objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronic Warfare</td>
<td>Suppress enemy EM spectrum capabilities to detect friendly assets while providing constant functionality despite enemy EM countermeasures</td>
</tr>
<tr>
<td>Capability</td>
<td></td>
</tr>
<tr>
<td>Defensive Capability</td>
<td>Maintain asset functionality against enemy attack and countermeasures</td>
</tr>
<tr>
<td>Offensive Capability</td>
<td>Provide offensive capabilities both within and beyond the battlespace</td>
</tr>
<tr>
<td>Power Projection</td>
<td>Project psychological power in the sphere of influence of the CSG to support peacekeeping missions, deter attack, and increase the prestige of the U.S.</td>
</tr>
<tr>
<td>Logistics</td>
<td>Replenish the consumables of the CSG anywhere in the world at appropriate intervals in combat and non-combat situations. The CSG shall be capable of operation for a reasonable period of time without re-supply</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expense Attribute</th>
<th>Notional portfolio strategic objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisition Cost</td>
<td>The cost of producing new constituent system</td>
</tr>
<tr>
<td>Influence Cost</td>
<td>The cost incurred to leverage a constituent system to participate in the CSG</td>
</tr>
<tr>
<td>Operation Cost</td>
<td>The annual cost of a constituent system</td>
</tr>
<tr>
<td>Schedule Cost</td>
<td>The perceived cost to portfolio designers of constituent system late entry into the portfolio</td>
</tr>
</tbody>
</table>

**Process 3: Capability Tree Stakeholder Value-Driven Design Formulation**

From the information gathered in the previous processes, the portfolio designer creates a notional structure of the capability tree by identifying specific program managers at each node. As a reminder, the manager at each node controls, or has expertise in, a performance capability identified in the higher level of the portfolio. The portfolio designer elicits a variety of information from each program manager of the capability tree and continues the process until all branches have terminated into the system-level attributes, or leaf nodes.

**Representational Outcome:** The capability tree for the CSG case study attests to the inherent complexity of portfolio analysis. For while the representation of the CSG has been simplified, the corresponding capability tree developed contains five branches (corresponding to the five portfolio-level performance attributes from Process 2) and, through up to four levels of decomposition, 91 distinct system-level performance attributes. For the sake of brevity, the entire capability tree has not been included in this paper; however, Figure 5 displays the tree with all but the system-level performance attributes included. It should be noted that some branches terminate after only the portfolio-level attributes, while other branches decompose capabilities through three nodes before reaching system-level attributes.
Figure 5. CSG Case Study Capability Tree Outline

Process 4: Epoch Characterization

All stakeholders related to the portfolio (portfolio decision makers, program managers, system operators, external stakeholders, etc.) are engaged to identify possible key future contexts that might impact success, and to characterize the uncertainty of each context. It is anticipated that a core set of contextual uncertainties will emerge from the stakeholders that may be parameterized by a common set of epoch variables. Any anticipated changes in stakeholder preferences (performance attribute weightings) are identified.

Representational Outcome: Seven epoch variables were identified from five major categories of uncertainty (technology levels, maintenance events, policy environment, SoS management abilities, and CSG threats). Table 4 displays the range of uncertainty represented by each epoch variable as well as units of measurement. The technology epoch variables most directly influence constituent system performance. The maintenance, policy and SoS management epoch variables most directly influence the cost functions. The threats epoch variables change the stakeholder preferences for portfolio attribute performance.
Table 4. Contextual Uncertainties of CSG Captured in Epoch Variables

<table>
<thead>
<tr>
<th>EV Category</th>
<th>Epoch Variable</th>
<th>[Range]</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>EV – Technology</td>
<td>Advanced Energy Weapons (AEW)</td>
<td>[0, 5, 50]</td>
<td>MW</td>
</tr>
<tr>
<td>EV – Technology</td>
<td>Unmanned Aerial Systems (UAS)</td>
<td>[2, 10, 50]</td>
<td>Berths</td>
</tr>
<tr>
<td>EV – Maintenance</td>
<td>Overhaul Event Costs</td>
<td>[0, 0.5e9, 2e9]</td>
<td>Billions $</td>
</tr>
<tr>
<td>EV – Policy</td>
<td>Budget</td>
<td>[80, 100, 150]</td>
<td>%</td>
</tr>
<tr>
<td>EV – SoS management</td>
<td>Cooperation Costs</td>
<td>[80, 100, 150]</td>
<td></td>
</tr>
<tr>
<td>EV – Threats</td>
<td>Enemy Threat</td>
<td>[Low, Med, High]</td>
<td>Level</td>
</tr>
<tr>
<td>EV – Threats</td>
<td>Asymmetric Threat</td>
<td>[Low, Med, High]</td>
<td>Level</td>
</tr>
</tbody>
</table>

For the initial case study presented in this paper, the technological epoch variables were excluded to simplify the analysis and focus the results on affordability considerations of the variance in portfolio cost attributes.

The five remaining epoch variables were enumerated independently and combined to produce a total of 243 potential epochs. For the sake of simplicity, five possible epochs that represent recognizable potential futures were extracted from this set in a “narrative” sampling approach. These five epochs and their epoch variable levels are provided in Table 5.

Table 5. Epoch Construction From Epoch Variables for Five Selected Epochs

<table>
<thead>
<tr>
<th>Epoch Names</th>
<th>Overhaul</th>
<th>Budget</th>
<th>Enemy</th>
<th>Asymmetric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (BL)</td>
<td>0</td>
<td>100</td>
<td>Low</td>
<td>Med</td>
</tr>
<tr>
<td>Small Navy (SN)</td>
<td>0</td>
<td>80</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>War on Terror (WoT)</td>
<td>0</td>
<td>100</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Major Conflict (MC)</td>
<td>0</td>
<td>150</td>
<td>High</td>
<td>Med</td>
</tr>
<tr>
<td>Peacekeeping (PK)</td>
<td>0.5e9</td>
<td>100</td>
<td>Med</td>
<td>Med</td>
</tr>
</tbody>
</table>

Process 5: System-Level Capability Assessment

Each potential constituent system must be evaluated for its performance in each system-level performance attribute. This “capability assessment” may take a variety of forms. In some cases, the bottom-level capability tree hierarchy managers may utilize performance models to computationally assess the potential constituent systems' performance. In other cases, however, such models may not exist and the bottom-level managers, or an appropriate subject matter expert, shall be engaged to assess system performance qualitatively. Each constituent system must be assessed for all performance attributes in each epoch.

Representational Outcome: The notional performance of the 19 potential constituent systems was assessed for each of the 91 system-level performance attributes in the five potential epochs. Publically available information was utilized to assign performance on a 0, 1, 3, 9 scale; 0 represented no performance, 1 was minimal performance, 3 was moderate performance, and 9 was performance sufficient to meet the desired portfolio capability. The CSG capability assessment results were reviewed for reasonableness by individuals familiar with naval systems.
Process 6: Design-Epoch-Era Tradespace Evaluation

A key high-level summary visualization is a tradespace with axes of Multi-Attribute Utility (MAU) versus Multi-Attribute Expense (MAE), as demonstrated by Wu (2014). However, unlike in traditional TSE, the points inside a portfolio tradespace do not represent single systems or programs, but rather represent unique combinations of assets determined by the portfolio investment strategy. The process flow of PLEEA depicted in Figure 2 highlights the process utilized to design portfolios, find portfolio MAU and MAE, evaluate the validity of the result, and create the data necessary for depicting a tradespace of viable alternatives.

Representational Outcome: Without the ability to elicit performance attribute aggregation models from each of the program-level managers, this research developed a series of six functions to evaluate bottom-level portfolio manager utility and cost attributes from the constituent systems in a portfolio. Each performance and cost attribute of the portfolio was assigned one of the six functions depending upon which one best represented the notional program-level manager value. Once the bottom-level performance and cost attributes had been calculated, the value was multiplied by the preference weighing of the program-level manager at the node in the next level of the portfolio hierarchy. This process was repeated for each level of the portfolio hierarchy until all branches were aggregated to the portfolio-designer. The portfolio-level preference weightings were then applied to find the ultimate utility and costs of the portfolio. Table 6 provides information on the approach and result for utility evaluation at each level of the capability tree.

Table 6. Approach to Derive Portfolio-Level Utility From System Performance Attributes Through the Capability Tree

<table>
<thead>
<tr>
<th>Portfolio Designer</th>
<th>2nd Level Program Manager</th>
<th>3rd Level Program Manager</th>
<th>System Performance Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach</td>
<td>Applied linear-weight sum aggregation to the utility from each 2nd level program manager</td>
<td>Applied linear-weight sum aggregation to the utility from each 3rd level program manager</td>
<td>Each system was evaluated on 0.1.3.9 scale for each performance attribute</td>
</tr>
<tr>
<td>Result</td>
<td>Portfolio utility on a 0 to 1 scale</td>
<td>2nd level program utility on a 0 to 1 scale</td>
<td>3rd level program utility on a 0 to 1 scale</td>
</tr>
</tbody>
</table>

The PLEEA method enumerated 53,018,336 possible portfolios and identified 524,160 portfolios which met the class constraints. This subset of portfolios was then evaluated according to the stakeholder preferences and constituent system performance of each epoch. Table 7 displays the number of valid portfolios, or portfolios which met the MAU and MAE constraints, and the percentage of the total potential designs that were found to be feasible. The Small Navy epoch is the most limiting epoch due to its 80% budget and 150% cooperation costs.

Table 7. Single-Epoch Tradespace Evaluation Summary

<table>
<thead>
<tr>
<th>Epoch</th>
<th>Valid Portfolios</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>156,750</td>
<td>29.9%</td>
</tr>
<tr>
<td>Small Navy</td>
<td>420</td>
<td>0.08%</td>
</tr>
<tr>
<td>War on Terror</td>
<td>156,750</td>
<td>29.9%</td>
</tr>
<tr>
<td>Major Conflict</td>
<td>492,400</td>
<td>93.9%</td>
</tr>
<tr>
<td>Peacekeeping</td>
<td>156,750</td>
<td>29.9%</td>
</tr>
</tbody>
</table>
**Process 7: Single Epoch Analysis**

The valid portfolios for each epoch may be plotted in a tradespace of MAU versus MAE. By exploring each single-epoch tradespace, a designer can identify which portfolio design options for the CSG perform particularly well for the context represented in that epoch, and what constituent systems Pareto efficient portfolios have in common.

**Representational Outcome:** The tradespace of viable CSG portfolios for the Baseline epoch is provided in Figure 6. The Pareto frontier of the Baseline epoch contains a total of 26 potential portfolio designs. The specifications for five example Pareto optimal portfolios have been provided. Figure 7 visually displays the composition of the promising portfolios and highlights constituent system investments. The constituent system types which differ between the promising portfolios were identified with Portfolio A as a reference. This portfolio comparison is intended to reveal to what degree the same constituent systems appear in various promising portfolios. Two portfolios which have the same types of constituent system, such as Portfolios A and C, are identified as having no different systems despite possessing varying numbers of each system type.

![Figure 6. Portfolio Affordability Tradespace (MAU v. MAE) and Pareto Efficient Portfolios for the Baseline Epoch](image)

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>MAU</th>
<th>MAE</th>
<th>Portfolio Composition</th>
</tr>
</thead>
</table>
| A         | 0.6536 | 0.0352 | Arleigh Burke Flight III  
1 Nimitz with Complement  
1 Los Angeles |
| B         | 0.7223 | 0.1487 | Arleigh Burke Flight I  
1 Arleigh Burke Flight II  
1 Arleigh Burke Flight III  
1 Nimitz with Complement  
1 Los Angeles |
| C         | 0.7548 | 0.1953 | Arleigh Burke Flight III  
1 Nimitz with Complement  
1 Los Angeles |
| D         | 0.7942 | 0.3072 | Arleigh Burke Flight I  
1 Arleigh Burke Flight II  
1 Arleigh Burke Flight III  
1 Nimitz with Complement  
1 Los Angeles |
| E         | 0.8358 | 0.6660 | NGCS Variant 6  
Arleigh B. Flight II Upgrade  
Arleigh Burke Flight III  
1 Nimitz with Complement  
1 Los Angeles |

![Figure 7. Constituent System Comparison for Promising Portfolios With Portfolio A as the Reference Portfolio](image)
**Process 8: Multi-Epoch Analysis**

A fundamental technique within EEA is the comparison of potential portfolios across multiple epochs. Following single-epoch portfolio analysis, a general understanding of performance of the SoS in each epoch has been established. By comparing the performance of each portfolio across multiple epochs, various metrics can be utilized to assess the design’s robustness against change and uncertainty, or how well a single portfolio can maintain its value across multiple epochs. The reader should consult Schaffner et al. (2013) for a detailed explanation of multi-epoch analysis.

**Representational Outcome:** A particularly useful concept to assess the performance of a promising portfolio design across multiple epochs, or to discover passive robust solutions, is the concept of Normalized Pareto Trace (NPT; Ross et al., 2009). An NPT score is assigned to each portfolio design of interest and describes the percentage of epochs in which that design constitutes a Pareto optimal point. A variant of NPT, the fuzzy NPT (fNPT), reveals the percentage of epochs for which a particular design is within a certain threshold factor of the Pareto front. The width of this threshold zone is defined by a K factor, where a K factor of zero indicates 0% fuzziness and is identical to the NPT metric. Figure 8 illustrates the concept of fNPT, and Table 8 displays the NPT and fNPT measures for the five portfolios identified in Figure 6.

![Diagram of fNPT metric](image)

**Figure 8. Illustration of the fNPT Metric, Where K Is the Threshold Width Applied Below the Pareto Frontier**

**Table 8. The NPT and fNPT Metrics for Five Promising CSG Portfolio Designs Over the Five Representative Epochs**

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>NPT</th>
<th>5% fNPT</th>
<th>10% fNPT</th>
<th>20% fNPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>C</td>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>E</td>
<td>0.4</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 8 displays that Portfolio A has an NPT value of one. This indicates that Portfolio A remains on the Pareto frontier in all epochs considered. From Figure 6, it is apparent that Portfolio A also represents the lowest expense and utility of any of the promising portfolio designs. Meanwhile the NPT for Portfolio C is 0.6, indicating that in 60% of the epochs, the design is Pareto optimal. However, Portfolio C reaches an fNPT value of one at a K value of 10%. This means that Portfolio C is within 10% of the Pareto frontier in terms of MAU and MAE for all epochs. These two portfolios (A and C) are the most passively robust to the uncertainties in the considered epochs. No other considered portfolio
beyond A or C has an fNPT value of one, even at a K value of 20%. This indicates that the
other promising portfolios are significantly removed from the Pareto front in at least one
ePOCH.

**Process 9: Era Construction**

Another useful technique in EEA involves the concept of evaluating portfolio designs
over an ordered set of epochs, called an era, to represent a potential lifecycle of the
portfolio. Era analysis enables the designer to understand how portfolio designs could
maintain their value through the uncertainty of a long run potential future. Disturbances and
degradation from an earlier epoch in the era may diminish the long term value of a particular
portfolio design with respect to another. Time-dependent concerns, such as cumulative
carrying costs and time to initial operating capabilities of various assets, can be considered
during era analysis. While such time-dependent factors were not considered as part of the
initial CSG case study, they are an anticipated area of future research.

*Representational Outcome:* Eras may be developed through stakeholder elicitation,
probabilistic modeling (e.g., Monte Carlo or Markov models), or through a narrative
approach to produce likely potential futures. Two eras were created for this case study. A
narrative approach was used to select the order and duration of the epochs to represent a
potential 30-year operational life for the carrier strike group:

**Era 1:** Baseline (5 yr), War on Terror (5 yr), Peacekeeping (10 yr), Baseline (3 yr),
Small Navy (7 yr)

**Era 2:** Peacekeeping (5 yr), Major Conflict (7 yr), Peacekeeping (10 yr), Small Navy
(5 yr), Baseline (3 yr)

**Process 10: Single-Era Analysis**

In single era analysis, the performance of portfolio designs is assessed over a
sequence of epochs as described in Process 9. Single-era analysis enables designers to
understand the time-ordered effects of the epochs on the portfolios. This allows for the
identification of portfolios which maintain utility and affordability throughout the sequence of
potential futures, and of those which may become unaffordable. The concepts of NPT and
fNPT may also be applied to an era to quantitatively measure how well a particular portfolio
design compares to the Pareto front solutions of each epoch in the era.

*Representational Outcome:* For the CSG portfolio design study, both the variance of
utility and expense of a potential portfolio over the lifecycle of the CSG are of interest to a
potential designer. Therefore, Figure 9 contains two subplots which display the change in
MAU and MAE over the five epochs for the five Pareto efficient portfolios in the Baseline
ePOCH. Tracing the trajectories reveals the emergent affordability, or unaffordability, of the
potential CSG portfolios.
Figure 9. (a) Expense Considerations for Era 1; (b) Utility Considerations for Era 1

The purpose of this initial case study was to explore the ability of PLEEA to support design for affordability. The case study was therefore constructed to focus on exogenous factors anticipated to impact expense attributes. As a result, the MAE of the portfolios exhibits significant variance over the era in Figure 9(a). However, there is relatively little variation in the MAU values between the selected portfolios through all epochs in Figure 9(b). The utility variation is small because the technology epoch variables, which most significantly influence constituent system performance, were not included in this initial study. Additionally, the class constraints of this analysis required all portfolios to have an aircraft carrier, a submarine, and at least one multi-mission capable surface combatant. The class constraints therefore provided a relatively high minimum utility of all valid portfolios which further reduces apparent variation. Finally, as illustrated in Figure 7, the promising portfolios primarily contained multiple units of the same constituent systems. As a result, the portfolios exhibit similar utility responses to the uncertainties modeled. Future research will include epoch variables which strongly impact capability performance attributes in order to more fully represent the design problem.

Process 11: Multi-Era Analysis

Multi-era portfolio analysis expands single-era analysis by identifying patterns of affordability (and unaffordability) as well as utility delivery that emerge from the path dependent development of portfolios through multiple contexts. This process enables a designer to develop metrics that characterize the affordability of potential portfolio designs across a variety of potential lifecycles. The reader should consult Schaffner et al. (2013) for a detailed explanation of multi-era analysis.

Discussion of the “Affordable” CSG Design Through PLEEEA

There are hundreds, if not thousands, of factors that influence the lifecycle design of a CSG portfolio. While it may not be possible to ever consider each of them, modern-day computing capabilities and systems engineering methodologies enable a designer to consider far more of these influence factors during conceptual design than was traditionally

---

1 This process was not demonstrated by the current study due to the representative nature of the analysis, but is described here for completeness.
possible. PLEEAA provided a method to consider 91 distinct CSG system-level performance attributes and relate their value back up to the overall portfolio utility. The performance attributes and expense attributes, along with the preferences of the relevant stakeholders, were defined in PLEEAA for five epochs to characterize contextual uncertainty of potential future states.

PLEEAA supported the identification of CSG designs which appear on the Pareto frontier of each epoch through single-epoch analyses. Furthermore, utilizing the metrics of NPT and fnPT in multi-epoch analysis, PLEEAA identified five promising portfolios which exhibit acceptable performance and affordability across a majority of the epochs considered. From Table 8, it can be seen that Portfolio A remains Pareto optimal through all epochs. Portfolio C, which may have been overlooked in traditional analysis as it does not appear Pareto optimal in two of the epochs, is the only other design which is Pareto dominant in all epochs when a 10% fuzziness factor is applied. Therefore, Portfolio C represents a passive robust solution that remains affordable against uncertainty, but is displaced in some cases from the Pareto frontier. Portfolios A and C are composed of identical types of constituent systems, simply with different numbers of these systems. This may indicate to a portfolio designer that the constituent systems in these portfolios provide promising value to potential future CSGs while maintaining affordability under potential uncertainty.

Finally, though single-era analysis, PLEEAA discerned which initially promising portfolio designs were likely to remain affordable, while maintaining their utility, over the lifecycle of the CSG. Figure 9 illustrates that while all portfolios maintain sufficient utility delivery, Portfolios B, D, and E become unaffordable in the Small Navy epoch. Upon investigation, these portfolios were found to exceed the operational costs allowable under the context of hypothetical Navy downsizing, while Portfolios A and C remained affordable. Single-era analysis of portfolios provides a mechanism to identify challenging lifecycle circumstances and robust solutions during conceptual design.

Discussion of PLEEAA Application to Portfolio Conceptual Design and Analysis

The conceptual design of systems of systems and portfolios presents a variety of challenges to traditional toolsets, including influence considerations, complexity of combination factors, and dynamically changing portfolio composition. However, with the prevalence of major acquisition Nunn-McCurdy breeches in recent years, the ability to evaluate the affordability of portfolios with respect to lifecycle uncertainty is desired to support acquisition decision makers. The PLEEAA method enables designers to enumerate and consider a greater number of potential portfolio designs than would be possible by current techniques. This is advantageous because it increases the ability of a designer to identify portfolios which display superior performance qualities such as affordability and robustness against exogenous uncertainty.

Tradespace exploration supports decision makers to identify portfolio compositions they may not have previously considered, to recognize macro-level trends in portfolio affordability, and to conduct micro-level tradeoff studies between portfolios in the area of desired performance. Epoch-Era Analysis provides acquisitions officers with new abilities to identify how potential designs may respond to a variety of anticipated contextual uncertainties, including the impact of simultaneously occurring uncertainties, in the lifecycle development and operation of the portfolio. The insight provided from these studies will support the selection of portfolios which may remain affordable over the entire lifecycle of the program, without the need to alter requirements or design in response to changing contexts. In implementation, the authors envision a network of system designers, each using
tradespace exploration to design their own systems, connected by a SoS and interacting through the PLEEA method to ensure that needed capabilities are provided in the portfolio at an acceptable cost with a desired level of uncertainty robustness.

Conclusion

The differences between capability portfolios of assets and individual systems necessitate specialized and distinct approaches for design and acquisition planning. Epoch-Era Analysis has shown promise in previous studies to enable the design of affordable systems that provide adequate utility while remaining under cost thresholds through a variety of potential contexts that may be experienced over the lifecycle of the system. Modern Portfolio Theory is a well-known financial tool that has been used for decades to select portfolios of investments which maximize utility subject to fixed uncertainty. This research leveraged EEA with elements of MPT to facilitate design for affordability of systems of systems with uncertain futures using a portfolio-based hierarchical perspective. The proposed method was used to explore 524,160 potential carrier strike group portfolio designs across five different epochs and two eras. This type of analysis may help decision makers to identify long-term acquisition strategies for affordable portfolios that are resilient against the types of contextual uncertainties that have negatively impacted recent DoD acquisition efforts.

References


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. RAND.


Acknowledgements

This material is based upon work supported by the Naval Postgraduate School Acquisition Research Program under Grant No. N00244-14-1-0018. The lead author would like to recognize the support of the NASA Aeronautics Scholarship Program under Training Grant No. NNX14AT14H.
An Analytic Model for DoD Divestments

Lisa K. Oakley-Bogdewic—is a principal staff economist at The MITRE Corporation’s McLean, VA, campus and has over 20 years of experience in various fields of economics in public and private sectors. Dr. Oakley-Bogdewic’s areas of interest are organizational structure and design, enterprise resource management and governance, strategic planning, policy design, process design, trade-space analysis, IT economics, performance management, and economic analysis to affect federal agency bottom lines. She holds a PhD from George Mason University in economics, with specialties in institutional economics, public finance, and labor economics. [loakley@mitre.org]

Emaan M. Osman—is a principal systems engineer at The MITRE Corporation’s McLean, VA, campus. Osman has over 20 years of experience in systems engineering, technical leadership, project management, and teaching. Osman’s technical focus includes systems-of-systems engineering; border security; mission command; Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C4ISR); acquisition; interoperability; architecture; experimentation; strategy development; systems analysis; trade space analysis; best value assessments; portfolio management; and assessments of IT efficiencies and effectiveness for government solutions. [eosman@mitre.org]

Abstract
This paper proposes an analytic model to improve Department of Defense (DoD) investment and divestment decisions. It builds from rules of thumb (RoTs) natural to government decision-makers and uses best practice models to improve the RoTs. It suggests three recommendations for the DoD: adopt the proposed model and address analytical barriers to inform RoT decisions, track investment vs. divestment decisions and Yes/No data points, and institute a Divestment Panel.

Introduction and Background
Organizations invest and divest resources to prepare for the future and respond to events or conditions in the relevant social, political, resource, or business environment (Scott, 2006). Successful preparation requires an effective and persistent process of management. Successful response requires a continuous and dynamic ability to offset threats and risks, or take advantage of opportunities. Both successful preparation and response require governance and management processes that focus decisions on the outcomes of the organization and improve the organization over time. Because resources are limited, prioritization of activities and deciding what to divest from is as important as making decisions on where to invest next, whether budgets are increasing or decreasing.

Deliberate Processes and Strategic Response
While commercial businesses deal with these challenges every day, this is difficult for the federal government for several reasons. The complexity, purpose (public goods), and

---

1 © 2015 The MITRE Corporation. All rights reserved. Approved for Public Release; Distribution Unlimited. Case Number 15-0971.
2 “Resources” can include capital, time, and any effort in activities.
scale present unique challenges to successful government preparation and response efforts (Wolf, 1979). First, the government sector in general, and the Department of Defense (DoD) in particular, is riddled with imperfect information and uncoordinated analytics. These often drive decision-makers to adopt rules of thumb (RoTs) to make what are usually suboptimal resource decisions (GAO, 2014). Second, best and possibly new practices are needed to simplify the logic and leverage natural biases already inherent to investment/divestment decision-making. Third, improvements in these factors would increase DoD leaders’ confidence of outcome impacts inherent in planning cycles, and to be more responsive to mission or fiscal disruptions.

Models of governance responsiveness are also needed to assess the ability of leaders to produce results, the accountability of leaders to do what is “right,” and the desire of stakeholders to make leaders more responsive in the public sector. This paper will look at best practices in the private sector and borrow applicable concepts from other complex risk-driven domains to derive and present a framework for government decision-makers to identify and govern the systematic divestment of low performing investments across the organization and free up funds for better operational and organizational choices. The proposed construct draws a corollary between typical commercial drivers and the DoD resource impact areas of Readiness, Modernization, and Force Structure. The construct then proposes to use and assess performance criteria of organizational and operational impact, performance viability of capabilities as core competencies of an organization, and the economic value and affordability of the investment set for the DoD.

**Human Decisions: “Imperfect” Resorts to Rules of Thumb**

Decisions to divest are often harder than decisions to invest. The culprit is often emotion or self-interest even when individuals believe they are doing the right thing. The challenge is compounded when the possession in question is providing some value, even if that value could be achieved more effectively elsewhere. For example, an organization decides to outsource an important non-core function to a specialty organization that can deliver it more effectively and cheaper than keeping it in-house. While this makes sense at an enterprise level, the internal providing unit meets the decision with resistance and reluctance driven by self-interest. In both cases, money spent or saved has direct impact on individuals involved in the function, jobs, or bonuses.

In the case of public money, the drivers may be less direct, but emotional nonetheless. Transparent governance is often cited as a long pole in the tent for effective decision-making, but add to this the notion that divestments are harder than investments, and the problem compounds (United Nations, 2003). Criteria-driven accountability and evidence- and performance-based management become crucial to preparing good investment and divestment decisions because they allow decisions to have a basis besides emotion. The absence of such criteria also results in organizations optimizing locally as opposed to the enterprise level; the end state is uncoordinated at best, and likely suboptimal for the enterprise. Without objective, consistently applied criteria, there is no repeatable, impartial way to assess performance of sub-organizations and sub-objectives and how they

---

4 There is a literature that discusses this fact even for “regular” stock markets, where there is a reward—of making money, or losing less of it—when a good divestment is made (Brown, 2013). See also Franklin Templeton Investments, 2013.
contribute to the success of the overall enterprise. Effective governance of these choices relies on symmetric and shared awareness of the impacts on enterprise goals and an understanding of the associated risks so informed resource decisions can be made. This paper proposes three criteria to enable the needed assessments:

1. **Organizational and operational impacts**: impact of decision on outcomes and enterprise goals
2. **Performance viability**: measurable performance of capabilities that are/are not core competencies for the mission and organizational outcomes
3. **Economic value**: investment costs and economic value of the investment/divestment set

**The Priority Order of Divestment May Not Be the Opposite of the Priority Order of Investment**

In the context of organizational decision-making, the list from which divestments are selected is not necessarily the opposite of the list of proposed investments. This is true for two main reasons:

- **Synergy and scale**: The whole is sometimes greater than the sum of its parts. Interdependencies may not be exposed until all the pieces are put together, and an attempt is made to break the whole apart or remove a part. Without tracked data supporting key types of criteria, the imperfect information and lack of causality in measurement causes pause when divestment proposals are made (Ebrahim & Rangan, 2010).

- **No financial meter or value proposition**: In public goods environments, such as Defense, the lack of a “bottom line” makes comparing the value of both investment and divestment choices difficult. For example, which is more important: a weapons platform, cyber security, the network, or force protection? Size of budgets becomes a meter and this does not promote divestment, nor does a notion of “affordability,” which is challenging to define and execute (see findings and progress at MORS, n.d.).

In addition to the uncertainties and omissions in decision making, cognitive biases may reinforce predisposed notions regarding investment or divestment options. Table 1 provides a summary of cognitive biases in the context of acquisition or divestment decision-making that are shown to result in rules of thumb (RoTs) that are used to make decisions in the absence of a performance- or evidence-based approach (Duhaime & Schwenk, 1985).

**Table 1. Cognitive Biases in Decision-Making**

<table>
<thead>
<tr>
<th>Investment Activity</th>
<th>Cognitive Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consideration of alternative acquisition</td>
<td>Reasoning by analogy: In the decision process leading to acquisition, decision-makers rely on analogies to simpler situations that may bear little similarity to their strategic problem. Illusion of control: In the evaluation of an acquisition, decision-makers overestimate the potential impact of their expertise on the acquired unit’s performance.</td>
</tr>
<tr>
<td>Management of the acquisition</td>
<td>Illusion of control: See above. Escalating commitment: Information on declining performance of the unit triggers rationalization and escalating commitment.</td>
</tr>
<tr>
<td>Consideration of divestment</td>
<td>Single outcome calculation: When divestment of a failing unit is finally considered, it quickly becomes the only course of action considered.</td>
</tr>
</tbody>
</table>
Without a data-driven understanding of the operational impacts and fiscal portfolio implications of investment sets, reasoning by analogy can turn perceived familiarity (“We have always done it this way.”) into a strategic error about projected outcomes and organizational and operational impacts. If not measured and tracked, the impacts generated by a particular investment set are easy to over-estimate with the illusion of control. This is especially true when emotional biases come into play (e.g., the idea that a broken program is fixable and is better than no program at all). Additionally, escalating commitment may arise, especially when a project is failing. Personal commitment to “save” the situation is natural for risk-averse thinkers. Risk averseness has been shown to increase with wealth, and should be relatively high for government fiscal stewards. These biases cloud the perception of performance viability, and true economic value of available options is not considered well enough to substantiate decisions. These all result in rules of thumb (RoTs) to help deal with complexity and ambiguity but yield poor choices (Center for the Study of Intelligence, 1999).

The Government Environment Arrives at Its Own “Rules of Thumb”

In the absence of shared awareness, legislatively required governance, constraints or negative feedback mechanisms, and evidence or performance driven decisions, public organizations generally end up with RoTs or traditions that appear to overcome common sense (see Table 2). The outcomes of low rigor can be inefficiencies and/or omissions that cause errors in decision-making in the aggregate, even though they may not have been made at the aggregate level.

5 While the income level of government workers may be disputable, and the monies they are working with are not their own, government employees are instructed to be highly regarded fiscal stewards, and we will stick with that theme in this paper. See Chu, Nie, & Zhang, 2014.

6 These rules of thumb were collected from MITRE subject matter experts in the area of acquisition and resource management.
### Table 2. Rules of Thumb: Common Government Divestment Approaches

<table>
<thead>
<tr>
<th>Common Government Rots</th>
<th>Resulting Divestment Behavior</th>
<th>Net Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Take money from fat programs or those that &quot;seem to have a lot.&quot;</td>
<td>&quot;Spend the money before it is taken.&quot;</td>
<td>When a budget is proportionally large, it becomes a target for redistribution and bill paying, even if the budget was needed to achieve desired outcomes.</td>
</tr>
<tr>
<td>2 Cut fairly.</td>
<td>Salami-slice (in current terms, &quot;sequester&quot;) cuts across the board, with a strict percentage target.</td>
<td>Impact and distribution of cuts are disproportionately detrimental to priorities.</td>
</tr>
<tr>
<td>3 Use trigger events (e.g., Nunn-McCurdy breaches) to mask as divestment.</td>
<td>Re-baseline programs and shift schedules to the right (slippage).</td>
<td>Money is simply spread over more years before a decision is finally made to kill a program or effort.</td>
</tr>
<tr>
<td>4 Penalize programs that measure performance.</td>
<td>Programs that effectively measure and report performance may be unfairly targeted and held to a different standard than those who do not.</td>
<td>Transparency is discouraged.</td>
</tr>
<tr>
<td>5 Pay for initiatives with &quot;urgency&quot; funds.</td>
<td>Gaps are resolved by taking money from unrelated programs.</td>
<td>Refusal to assess baseline for cuts related to gaps infers no business case is formulated—this can become a habit.</td>
</tr>
<tr>
<td>6 Congressional disregard or favoritism</td>
<td>This drives, or saves, many divestment decisions for all military departments alike.</td>
<td>Decisions are made based on what can be approved, not what is needed.</td>
</tr>
</tbody>
</table>

### Business Best Practices to Deal With Cognitive Biases

Because risks and returns are involved in divestment choices, the private sector has developed best practices (Mankins, Harding, & Weddigen, 2008). Table 3 compares best business practices against the DoD’s common government practice.
### Table 3. Divestment Commercial Best Practices vs. Common Government Practices

<table>
<thead>
<tr>
<th>Best Commercial Business Practice</th>
<th>Cognitive Bias</th>
<th>Common Government Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Dedicate a team to divestment full-time, just as you do with acquisitions.</td>
<td>Reasoning by Analogy</td>
<td>DoD has thousands of people that work on strategic and investment planning. Not many focus on divestment, but closing funding gaps.</td>
</tr>
<tr>
<td>2 Make sure you can clearly articulate how the deal will benefit the buyer and how you will motivate the unit’s employees to stay until the deal is done.</td>
<td>Illusion of Control; Escalating Commitment</td>
<td>Understanding costs of transition and divestment is a captured, but under-appreciated set of costs in the DoD because transition takes a very long time with many changes in leadership</td>
</tr>
<tr>
<td>3 Work through the details of the de-integration process before you divest.</td>
<td>Single Outcome Calculation</td>
<td>DoD’s duplicate costs for capabilities or services are on the books, yet may change names over the transition time, and are difficult to track. Limited insight into interdependencies</td>
</tr>
<tr>
<td>4 Establish objective criteria for determining divestment candidates—don’t panic and sell for a song in bad times</td>
<td></td>
<td>While DoD demands auditability, it does not emphasize accountability for losses, sunk costs, and returns on investment (Under Secretary of Defense [Comptroller]/Chief Financial Officer, 2013)</td>
</tr>
</tbody>
</table>

### Towards a Solution

The model presented here proposes that decisions should be made based on three key factors when selecting investments or divestments (Campbell & Whitehead, 2014):

- **Strategic value:** the criticality of the investment to operations of the enterprise
- **Performance:** potential to improve the business or create synergy with other businesses
- **Economic value:** net present value (NPV), capital flows (capital required and lifecycle efficiencies), and external or secondary effects

### The Proposed Framework and Criteria

Figure 1 below diagrams the investment/divestment logic for a commercial business (Suozzo, 2001). There are three general cases in Figure 1: (1) “obvious” buy/sell cases, (2) situations where the candidate is not strategic, and (3) cases where the candidate is strategic, but the required competencies or cost advantages are not necessarily evident. These latter two cases are most interesting for investment houses, but they are all interesting for government, because even “obvious” buys are an opportunity cost of another.

---

7 As repeatedly witnessed by the authors.
8 DoD architectures are listed in the Defense Acquisition Registry System (DARS): www.dodenterprisearchitecture.org/exhibits/Pages/DoDArchitectureREgistrySystem(DARS).aspx
choice, so they must not only answer “Yes” to all three criteria, but the choice must also surpass the holistic value of competitors for the resources. Figure 1 makes commercial business sense.

![Diagram](image)

**Figure 1. Three Criteria for Investment/Divestment for Commercial Business**

But what about self-interest and emotion? The “unmaking” of a product line? The transition of support to potential partners? These factors alter the speed and cost of transition, which are guided efficiently by negative feedback mechanisms that ensure survival (Hardin, 1968).9

How might this apply to the DoD? For example, in response to wars, a DoD resource manager may choose to invest because the capability is strategic and high performing, but refuse to consider the high budget share and opportunity cost, or the risks that are unknown, or because the impacts to the defense industrial base are unacceptable. We end up with an array of buried or obfuscated investments that would benefit from an independent divestment panel equivalent. Figure 2 presents a proposed mapping of the business model to the DoD’s key drivers. Not only are the DoD resource concepts of Readiness, Modernization, and Force Structure attributes that the DoD seeks, but they are also Resource areas reflected in the budget (Trunkey, 2013).

---

9 This article is actually a testimony to what happens when decision-makers choose to act in their own self-interest and are not guided by higher order feedback mechanisms, such as the over-grazing of public lands or the pollution of air.
Figure 2. Translation of Valuation Criteria to DoD Resource Areas

Figure 3 shows 70 years of DoD trends approximating the Congressional Budget Office’s (CBO’s) budgetary definitions of Readiness (all Operations & Maintenance [O&M], except Healthcare, Revolving funds, and Civilians), Modernization (Research, Development, Test, & Evaluation [RDT&E], Procurement, Military Construction, Science, & Technology, and Weapons Acquisition), and Force Structure (Military Personnel, Military Health [O&M], Family Housing).

Figure 3. Seventy Years of Spending Trends: Impact (Readiness), Performance (Modernization), and Economies (Force Structure)

At any given time, resources are developing, equipping, organizing, training, sustaining, and manning the force, and many dollars may “cross-over” and affect other resource pools. The high manpower of the 1950s shows the human warfighting power needed in the Korean War, the high Readiness costs in the wars since 9/11 shows the force is highly equipped for operations, and the increase in modernization of the 1980s shows the DoD’s strategy to the Cold War.

These appear to be logical given the environments, but there is no objective way to assess the relative goodness of the decisions here. Additionally, a criticism across the three DoD processes—Requirements, Acquisition, and Budgeting—are that the criteria used for decision-making are not defined or well-coordinated (Defense Business Board, 2012):
The processes do not have contemporaneous objectives (some are farther out than others).

They track the resource pools differently: joint capability areas, program executive office families, Budget Activities, Program Elements, Appropriations, etc.

DoD goal achievement is not measured, tracked, or aligned to resource pools to measure any kind of efficacy of planning or delivery.

While Figure 3 reflects the striking of an appropriate balance, it is impossible to determine within the current construct what balances have been achieved (Sledge, 2010). Figure 4 provides a summary of the model developed in this paper and shows how an analyst can arrive at answering “Yes” or “No” to the evaluation criteria proposed.

Figure 4. Proposed Investment/Divestment Framework for the DoD

Figure 5 borrows the logic from Figure 2 and puts the concepts in DoD terms to propose a set of criteria to assess such balance. This modified framework applied to the DoD should enable the use of a common language and valuation approach so that decisions could be made on common criteria of readiness (operational effectiveness), modernization (capability performance promise), and force structure efficacy (sustainable cost). Every investment set can be assessed against these three concepts with a “Yes” or “No” answer. We assume that all investments are compared from an end-to-end basis.
Proposed Criteria

Criteria #1: Strategic Value, Impact Readiness

Congress uses DoD-provided quarterly readiness assessments to determine resourcing requirements of the military (Trunkey, 2013). Therefore, any resource-consuming activity that occurs within the DoD should promote or support readiness either directly or indirectly. If a relationship between an activity and readiness cannot be demonstrated or articulated, then the determination should be “No” to this criterion. All investment sets could be placed in a priority order of Readiness contributions, and a cut line could be established.

---

10 Much of this material was taken from Trunkey, 2013, and Osman, Wilk, and Oakley-Bogdewic, 2013.
Much effort is spent assessing how to measure how the DoD resources readiness, in addition to measuring Readiness levels.\(^{11}\) The assessments are for units ("unit readiness") and organizations within the military departments, which is not necessarily about an investment but the ability of the outputs supported to fill the critical operational logic presumed for the mission. Note that it is possible for an investment to have a “No” determination for this criterion and still be a viable investment choice. For example, activities that do not have a visible impact on the effectiveness of units but that allow them to execute their missions more efficiently or economically could be viable investment candidates.

To meet Criteria 1, the investment set must meet critical equipping, training, and operations needs for current or future missions of critical Combatant Command forces or be vital to the accomplishment of a QDR Goal.\(^{12}\) Outcome indicators are needed for each investment set to track expectations of impact.

**Criteria #2: Core Competency, High Performance, or Modernized**

To satisfy Criteria #2, investment sets should be able to answer one or more of the following:

- **Is it a Core Competency?** (Chairman of the Joint Chiefs of Staff [CJCS], 2012a, p. 2).\(^{13}\) The DoD has nine Joint Capability Areas (JCAs), each with Capability Based Analyses and current validated requirements. Each has a Functional Control Board that oversees and rationalizes its area’s

---

\(^{11}\) Recent examples include the following:
From the CBO, see (1) CBO, 2013, p. 13; (2) CBO, 2011, pp. 3–4 & 8: This report separates Mission-Related and Infrastructure Related readiness spending; (3) Trunkey, 2013, pp. 1 & 17: Trunkey uses Personnel, Equipment, Supply, and Training in his definition (p. 17), as he states, “DoD spends about $350 billion to further the readiness of its forces for current and future military operations. This includes attracting, retaining, educating, and training top quality military personnel; keeping equipment well maintained; and providing the food, fuel and other material needed to support operations” (p. 1).
From the GAO, see (1) GAOb, 2013, Table 1; (2) GAO/NSIAD-95-29 (GAO, n.d.).
From the DoD, see DoD, 2013, Annexes A–M: Each of these functions can potentially be tied to one of the three concepts: A. SPACE; B. CYBER; C. CIVILIAN PERSONNEL; D. MILITARY PERSONNEL; E. READINESS; F. MUNITIONS; G. ACTIVE COMPONENT/RESERVE COMPONENT (AC/RC); H. NAVAL PRESENCE; I. STRATEGIC DEFENSE AND DETERRENT; J. SPECIAL OPERATIONS FORCES (SOF); K. INTELLIGENCE, SURVEILLANCE, AND RECONNAISSANCE (ISR); L. ENERGY; M. INDUSTRIAL BASE & THE OFFICE OF ECONOMIC ADJUSTMENT.
From the CSBA, see Harrison, 2012, p. 16: Harrison stated, “Readiness funding includes the O&M budget activities for Operations Forces …, Mobilization …, and Training and Recruiting.”

\(^{12}\) Criticality for COCOM forces is planned for by each service. See, for example, U.S. Army, n.d., Chapter 7. The QDR lays out the foundational force structure required from each service, which recently, was tied to the FY15 budget submission funding the force structures. See also [http://www.defense.gov/releases/release.aspx?releaseid=16567](http://www.defense.gov/releases/release.aspx?releaseid=16567) and [http://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2015/fy2015_Budget_Request.pdf](http://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2015/fy2015_Budget_Request.pdf).

\(^{13}\) While it is true that the Joint Staff keeps track of all requirements in the DoD, not all investments have formal requirements: “The responsibilities of the JROC over ‘joint military requirements’ include both joint requirements and single DOD Component requirements which makeup the entirety of the capabilities of the joint force and enable the DOD core mission areas” (CJCS, 2012 p. GL-6).
requirements (CJCS, 2012a, p. 2; CJCS, 2012b, Enclosure G). The DoD also conducts a Quadrennial Defense Review (QDR) every four years, and key goals are established with this review.

- *Does it deliver High Performance?* The Joint Capability Integration Development System (JCIDS) Manual provides a guide to parameterizing performance, which is a full array of metrics to gauge performance of a military system’s actual and projected performance: Capability Performance (each JCA has defined performance attributes), Force Protection, Survivability, Sustainment (which includes Reliability and operating and sustainment costs), Net-Readiness, Training, and Energy (CJCS, 2012b). Even if an investment set is not in the JCIDS process, these serve as holistic and common criteria.

- *Is the Modernization contribution needed?* The age distribution of investments in the capability area within which the investment set lies suggests the importance of modernization to the capability area, and the degree to which the funds supporting the investment set in question are value-added. A third metric would assess the degree of modernization existent in the capability area and the expectation of the investment set’s contribution to this maturity. This would include market research on the DoD’s need to have a leading edge capability.

**Criteria #3: Economically Sound or Force Structure Supportable**

Investors solely seeking a return on investment (ROI) have beliefs about factors such as strategic value and performance for their financial needs. These beliefs lead to levels of confidence in investing, which in turn drive preferences over how much to change from, or divest from, their status quo portfolio. The preferences translate into behaviors such as larger numbers of trades or wider stock diversification. Investors using more highly informed analytics traded more (and therefore divested more), had more diverse portfolios, and had higher returns (Hoffman et al., 2010). With the general objective to manage the risk profile of their consumption stream, investors of all risk types (from low to high: retirement, financial diversification, capital growth, hobby, speculation) have various tendencies to “stick to the status quo” and not divest. 

**Leveraging Other Fields: Social Return on Investment & Insurance Models**

This evidence encourages the proposed DoD divestment framework to include an economic soundness calculation that shows savings or efficiencies for the budget environment (Moore, 2009) and consider a wider measure of Social Return on Investment (SROI). Not only will the SROI calculation consider the value of strategic impact and performance, the SROI concept takes into account two new variables—secondary effects.

---


15 From Hoffmann et al. (2010, p. 16): “Status quo bias is strong.” Mitchell et al. (2006) provide evidence that 80% of participants in 401(k) accounts initiate no trades in a two-year period, and an additional 11% make only one trade. Therefore, few investors in their sample rebalance. Similarly, Ameriks and Zeldes (2004) find that 50% of the investors in their sample do not rebalance over a nine-year period.
and deadweight (Rauscher, Schober, & Miller, 2012)—that make divestment decisions more difficult (Rauscher et al., 2012, p. 6):

- **Secondary effects** include impacts on other portfolios, including intended externalities and unintended consequences (e.g., the effects on other portfolios or the defense industrial base as a result of terminating a program). Enterprise architectures and cross portfolio management should capture and account for these.

- **Deadweight** is a placeholder for the levels of productivity or outcome changes that would have happened anyway, without the intervention (e.g., technology obsolescence or a politically driven decision to withdraw troops). The intent is to factor out double counting, especially if more than one or a complex intervention is being assessed.

Based on models of social investment, social entrepreneurship, and venture philanthropy, the SROI model captures varied types of impacts and outcomes. It focuses on an investment “intervention” at the enterprise, program, or project level. In addition to mission effects, their SROI calculation includes how the intervention affects how the organization functions as well—activity efficiencies—through the size of secondary and deadweight effects, as shown in Figure 6.

Time and portfolio averages taken from the insurance industry as variables that impact actuarial calculations are also worth considering here. For example, as the insurer of national security, the DoD’s needs, risks, and cost patterns evolve with threat cohorts, weapons, and technology generations over time. How these factors may change and complicate lifecycle calculations for costs and returns cannot be overlooked (Lebar, 2012; Wadsworth & Woodley, 2013). These added concepts parameterize factors that usually “blur” or are omitted from the standard DoD return on investment calculation, making divestment decisions more difficult. Figure 6 summarizes the logic model behind the SROI concept.
Note. Terms and definitions used in Figure 6:

**Context:** Economic, political, and social circumstances
**Income:** Target group needs; for the DoD, this can be a mission outcome
**Input:** Financial and personnel resources available
**Structure:** Legal and financial characteristics of sponsor
**Concept:** Roles, responsibilities, due dates of targets
**Process:** Targets linked to activities
**Output:** Directly provided contributions of program/activity to achieve desired impact(s)
**Outcome(s):** Effects or desired conditions of target group after activities are completed
**Impact:** Overall effects that are logically, theoretically, or empirically substantiated
**Initial:** Time for initial intervention impacts to be realized
**Interim:** Time for secondary effects to be observed and quantified/qualified
**Long-Term:** Time for relationships between impacts and deadweight to be delineated

**Figure 6. Logic Model Behind SROI Concept**
(Rauscher et al., 2012)

In addition to Secondary Effects and Deadweight, our model leverages two factors from insurance modeling as important considerations (Lebar, 2012; Wadsworth & Woodley, 2013):

- First, there is a time element critical to returns on the portfolio. Time is important because of the timed targets set in the plan and because performance of incoming investments in the intervention may be masked by portfolio-wide calculations of ROI. When older investments retire, the ROI for the portfolio may rise or fall dramatically.
- Second, risk categorization errors may occur in NPV calculations. For example, models may rate a threat (or policy) being abated by the intervention as higher or lower than it should be. Not only will the error apply to the investments in the intervention, but it may also affect secondary effects or deadweight.
What the Model Tells Us

While the detail in each criterion is potentially exhaustive, holistically the information derived from Figure 4 should address some of the biases discussed in Section 2.0 and enable improved decisions. Table 4 threads the constructs discussed in this paper. The generally omitted criterion in the first column leads to the cognitive bias usually relied upon in the second column. This leads to the DoD RoTs for divestment in the third column, to commercial best practices recommended to avoid mistakes typically made with these biases. These tie to the simple Yes/No criteria presented above.

Table 4. Tying It All Together

<table>
<thead>
<tr>
<th>Omitted Criterion</th>
<th>Cognitive Bias</th>
<th>DoD RoTs for Divestment</th>
<th>Best Practices</th>
<th>Model Y/N Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational &amp; Operational Impacts</td>
<td>Reasoning by Analogy</td>
<td>- Take from those who have a lot&lt;br&gt;- Cut fairly (salsami)</td>
<td>Dedicate a (small) full-time team to Divestment planning</td>
<td>Importance: Is the investment set Mission Critical?</td>
</tr>
<tr>
<td>Performance Viability</td>
<td>- Illusion of Control&lt;br&gt;- Escalating Commitment</td>
<td>- Use trigger events to re-baseline&lt;br&gt;- Penalize the transparent</td>
<td>Establish Core/Non-core; Measure primary, secondary, deadweight values; Plan details of divestment transition</td>
<td>Confidence: Do we have confidence in the value proposition?</td>
</tr>
<tr>
<td>Economic Value</td>
<td>Single outcome calculation</td>
<td>- Pay for new w/ unrelated funds&lt;br&gt;- Congressional favor</td>
<td>Account for the de-integration costs; Track variables to calculate SROI, including secondary, deadweight effects</td>
<td>Fidelity: Can we account for both value contribution and total cost factors for the portfolio baseline and the intervention?</td>
</tr>
</tbody>
</table>

Table 5 summarizes the options across three factors and Investment versus Divestment Decisions.

Table 5. Relations Between Factor Acceptance and Invest/Divest Decisions

<table>
<thead>
<tr>
<th>Criteria Area</th>
<th>Factor</th>
<th>Discriminator</th>
<th>Acceptance</th>
<th>Invest</th>
<th>Invest</th>
<th>Invest</th>
<th>Invest</th>
<th>Divest</th>
<th>Divest</th>
<th>Divest</th>
<th>Divest</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Strategic Value or Readiness</td>
<td>Operations Logic Critical</td>
<td>Investment sets meets critical equipment, training, and operations needs for current or future missions of critical COCOM forces or vital to accomplishment of a CDR goal.</td>
<td>Mission Critical?</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>High Performance or Modernization</td>
<td>Core competency?</td>
<td>Validated Requirement</td>
<td>Included on Unified Command Integrated Priority list?</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>High performance?</td>
<td>Performance in array of NPP metrics*</td>
<td>Average above Threshold?</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Modernization impact?</td>
<td>Impact on Balance of investment types</td>
<td>e.g., profile defendable w/ Mod needs</td>
<td>Y</td>
<td>otherwise</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>either</td>
<td>either</td>
</tr>
<tr>
<td></td>
<td>Cost Effective</td>
<td>Cost is supportable, LPO outsourced, or is income</td>
<td>Advantageous NPI?</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>either</td>
<td>either</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>ECONOMICALLY Sound or Force Structure Sustainable</td>
<td>Secondary Effects+ insured</td>
<td>Intended externalities, unintended consequences</td>
<td>Are SE measurable? Do they add/subtract as NPI is worthwhile?</td>
<td>either</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>either</td>
<td>either</td>
</tr>
</tbody>
</table>
The following are key observations from Table 5:

1. The economic factors drive the decisions.
   - **Cost Effectiveness** drives investment except when mission criticality and performance are high, and secondary benefits are high and the deadweight is minimal.
   - **Deadweight**—Other factors actually drive the outcome, not the investment—uniquely determines investment and must be accounted for.
   - Only in one case are the **Secondary Effects** immaterial to investment choices, where everything else is affirmative (the first column under Invest).

2. High performance or modernization impact worked with the economic variables. The **Impact on Balance of Investment Types** is also nearly collinear with invest/divest decisions.

3. **Mission Criticality** may not be necessary for investment selection (e.g., when investments drive costs down while minimally impacting missions).

These conclusions are not natural products of the traditional RoTs discussed in the Introduction and Background and further demonstrate how uninformed RoTs lead to suboptimal decisions.

Figure 7 shows some well-known DoD investment and divestment choices overlaid on the proposed model. For example, an obvious investment area is Classified Networks since it easily meets all three criteria presented by model. The decision to divest of the next generation of amphibious assault vehicle is also supported by the model since while useful, it does not materially appear to affect the DoD’s ability to execute missions (not mission critical), and, crucially, it presented limited force structure supportability. An example of a non-mission critical/non-strategic recommended investment area is Defense Finance and Accounting Services (DFAS) because it is important for efficiently operating the enterprise while minimally impacting actual execution of operational missions.
Recommendations

From a very wide discussion of rules of thumb in environments with imperfect information to SROI models to DoD criteria and data, we have drawn together an understanding of why persistent investment and divestment management data and processes can overcome predictable mistakes in resource decisions. We now present three basic recommendations:

1. Adopt the model presented in this paper and address the analytical barriers to allow decision-makers to answer the three basic questions presented in the model.
2. Track the logic of Yes/No data points as a feedback mechanism to decision-makers.

Recommendation 1: Adopt the model presented above to enable shared awareness and improved analytics so decision-makers are no longer forced to rely on uniformed RoTs to make decisions.

Start small. Use the model for key program choices, and then for capability area choices (such as a program executive office), and then for portfolio choices. Understand the types of data needed to answer the questions in Figure 4. Begin using data from the DoD Requirements, Acquisition, and Budgeting processes to answer questions previously answered by rules of thumb. Enable the population and use of the SROI model, which incorporates data from all three criteria sets. The use of deliberate criteria could lead to the reward of auditability and the delivery of better results.
Recommendation 2: Track the logic of Yes and No investment data points—at least at a capability level—so that the logic is captured, results are transparent, and the organization improves and learns from successes and mistakes. This will perpetuate responsive governance.

Create a database of Invest/Divest decisions, traceable to the three criteria, and to the actual results produced over time. This traceability is difficult, as investments get joined and split, and the investment decisions made at the time are often altered before investments are actually implemented. The intent can be tracked, however, and the investment and divestment strategies can be logically understood. This should help stage migration plans so that Divestments become part of the usual process. Make decisions using the proposed framework and encourage the use of analytics, tracked to productive implications.

Recommendation 3: Adopt an industry best practice of divestment panels as a way to promote the best use of resources instead of relying on rules of thumb as reliable predictors.

The DoD should adopt the framework described in the earlier section, Human Decisions: “Imperfect” Resorts to Rules of Thumb, as a method for determining investments with simple and meaningful criteria while at the same time employing a divestment panel as a portion of their means of governance.

It should also establish standing divestment panels comprised of non-advocates who report their results annually to senior Component leadership. Investment panels exist at many levels in the DoD, but per Title 10, the Secretary of Defense and the Military Secretaries have the final say on divestment.

The DoD has the ability to divest, for example, when politics calls for “peace dividends.” In response to the 2011 Budget Control Act, the DoD was able to divest from $500 billion per year (7%–10% of totals, varied across organizations). The DoD reduced Force Structure and also found “efficiencies” under the oversight of Secretary Gates (see the Defense Strategic Guidance [DoD, 2012]; DoD, 2010). Cuts can be taken in Strategic/Readiness endstrength equipping and training, Performance/Modernization upgrades to weapon and systems portfolios, or Economic/Force Structure areas of manpower-related investments. Savings take time to realize and are difficult to account for when continued operations and innovation are still taking place (Pellerin, 2013).

The DoD also has the Issue Paper process supporting RMD-700, which allows community leaders across the DoD to recommend shifts of resources before the budget is finalized. Money moved in this process is usually a small fraction of total spending and is considered fallout from the annual budgeting process (Huo, 2011; USD[C], 2013). Both types of formal divestment processes are normal parts of the current bureaucracy, but do not necessarily illustrate a governance process relying on an informed thought construct to foster and maintain a forward-looking, innovative organization.

---

\[16\] In the Fiscal Year (FY) 2014 President Budget Submission, the USD(C; 2013) claimed, “RMD 700 identifies a limited number of DoD-wide performance goal priorities” and suggests that “DoD Component-specific budget justification material should be consistent with” these goals.
In a budget-scarce environment, divestment strategies such as those recommended in this paper could prove to be useful for justification and selection of the “keeper” investment sets. The impact of a divestment panel could be measured. If the DoD captures and tracks distinct data on strategic impact, performance, and economic efficiencies, the DoD and Congress could reward good results with leadership recognition and continued fiscal support to sustain the positive pattern.

Summary and Conclusions

This paper outlined a logical investment/divestment choice structure that corrects old RoTs that predictably lead to suboptimal choices. It relies on decision-makers having the capability and reward structure to use such a structure in a complex environment. It assumes that the DoD institution would be motivated by Congress to hold decision-makers accountable, and in turn, these leaders would be rewarded for being responsive and productive in their choices. In this new frame, divestments would be seen as strategic opportunities, and the reward structure would incentivize accountability and measurable outcomes. The following are two recommended next steps:

- **Design a tool for Preparation**: Gather data to answer Yes/No questions; program tool to produce Invest/Divest choices; alter model’s basis with events.
- **Use for effective Response**: Use analytics to increase capability of leaders; link analytics to “accountability data”; and effectively reward organizations and their leaders.

References


Congressional Budget Office (CBO). (2011, April). Linking the readiness of the armed forces to DoD’s operations and maintenance spending.


Portfolio Acquisition—How the DoD Can Leverage the Commercial Product Line Model¹

Peter Modigliani—is the Acquisition Innovation Area lead at The MITRE Corporation. He supports DoD acquisition and CIO executives’ strategic initiatives in Agile, cyber, IT, and services acquisition. He manages a research portfolio to foster innovative acquisition solutions. Previously, as an assistant vice president with Alion, he supported the Air Force Acquisition Executive on C4ISR systems. As an Air Force program manager, he developed strategies for billion dollar acquisitions. Modigliani holds a BS in industrial engineering from the Rochester Institute of Technology and an MBA from Boston College. He is DAWIA Certified Level III in program management. [pmodigliani@mitre.org]

Abstract
The Department of Defense (DoD) can foster dynamic, efficient, and innovative solutions for tomorrow’s warfighter by structuring acquisition portfolios that deliver an integrated suite of capabilities. Such portfolios would permit execution of many core acquisition elements and processes at a level above the individual program to enable enterprise management, economies of scale, and faster capability deliveries. While large DoD programs navigate the acquisition life cycle individually, large commercial businesses manage integrated product lines for items ranging from automobiles and personal electronics to software and health services. The portfolio framework proposed in this paper establishes broader entities that involve an active government and industry community throughout the acquisition life cycle. Portfolios would scope programs and increments from high-priority requirements, mature technologies, and rigorous analyses covering a comprehensive mission area. Portfolio strategies, roadmaps, and architectures would guide development of a suite of smaller programs, allocating budgets, personnel, and other resources dynamically to the highest priority efforts. Reorganizing from a product-based model to a portfolio model would enable more successful and faster delivery of integrated mission capabilities.

Introduction

Challenges of the Program-Centric Acquisition Model
In today’s defense acquisition system, each program navigates the acquisition life cycle individually. This results in an acquisition enterprise that leads to stove-piped solutions, long acquisition cycles, and a highly inefficient use of resources. Initial conceptual requirements drive program scope and budgets, yet often inappropriately constrain the solution space for long-term programs that develop major systems. The lengthy congressional approval process for new start programs contributes to setting a high bar up front to DoD exploration of new solutions.

Developing systems individually makes it extremely challenging to deliver the integrated, net-centric systems and services required for the DoD’s complex and dynamic operations. Acquisition programs design, develop, test, and produce isolated systems that must meet a defined set of requirements within an allocated budget. Analyses of alternatives (AoAs) occur at the program level, with minimal consideration of enterprise performance,
costs, or risks. Each program must conduct its own research and development (R&D) to mature its critical technologies in order to begin development (see Figure 1).

Figure 1. Program Silos in the Current Acquisition Framework

Guiding large systems through the acquisition life cycle over a period of 10–20 years has proven inefficient and ineffective as technologies, operations, and budgets change. Selecting a development contractor alone takes a year or longer, and in the process programs often lose critical insights that could be gained from subsystem prototypes and preliminary designs. As other nations rapidly adopt commercial technologies and exploit global networks, the DoD’s technological advantage confronts greater risk.

In the years 2001‒2011, the DoD spent over $46 billion on Major Defense Acquisition Programs (MDAPs) that were ultimately canceled (Harrison, 2011). A major contributing factor common to these failures is that the programs tried to do too much at once: they used a big-bang approach to develop and integrate a wide array of technologies to meet all envisioned requirements. For example, the Army’s Future Combat System (FCS) attempted to develop a dozen classes of ground systems, unmanned aerial and ground vehicles, and an integrated network as a single MDAP; FCS was cancelled after spending $18 billion. The Air Force’s Expeditionary Combat Support System (ECSS) sought to replace over 250 legacy logistics information technology (IT) systems with a single new system and invested $1.1 billion and nearly a decade of effort before program cancellation (Levin, 2014).

The DoD’s acquisition budget has been reduced by tens of billions of dollars annually from the levels of the previous decade. The DoD’s fiscal year (FY)2015 research and procurement budgets alone have declined by 21% and 29%, respectively since FY2010 (Weisgerber, 2014). In an era of continued global threats, the DoD could lose its technological edge unless it takes bold steps to structure and streamline the acquisition framework to deliver capabilities to the warfighter more effectively (Kendall, 2014). To accomplish this, the DoD must leverage the structure and methods of large commercial enterprises, including auto manufacturers, consumer electronics companies, and professional services firms, all of which use product lines to obtain the greatest benefits from their investments.

Commercial Product Lines

Commercial firms use an approach that evolves a product to its ultimate capabilities on the basis of mature technologies and available resources. This
allows only the product features and capabilities achievable with available resources in the initial development. Further product enhancements are planned for subsequent development efforts when technologies are proven to be mature and other resources are available. (Walker, 2013)

Many large corporations organize along product lines to leverage economies of scale and react swiftly to emerging trends and changes in consumer demands. For example, Apple revolutionized consumer electronics because it did not simply develop products that outperformed others in the marketplace but focused on delivering a full integrated user experience across products and services. Toyota designs, develops, and produces its cars, trucks, and SUVs by leveraging technology innovations across all of its models.

With many Fortune 500 companies facing strong challenges from emerging startups, executives are aggressively breaking down corporate silos and reengineering operations to pursue innovative solutions. Leading companies embrace “design thinking” that prompts them to observe market nuances, experiment with many options, and rapidly prototype ideas to bring the best ones to reality. They maintain strategic variety, to include creating portfolios of new strategic options, building a magnet for great ideas, and minimizing the cost of experimentation (Hamel, 2012).

Companies designate product line managers to maximize revenue and profit from the company’s investments and executives grant these managers significant latitude to shape the products they manage. This includes marketing, developing new products, forming corporate partnerships, and conducting R&D. The success of a product line depends on the company’s ability to track the market closely and react faster than the competition to emerging trends, technology advancements, and changes in consumer tastes. The success of this strategy, in turn, stems from aligning each product line manager’s responsibilities with accountability: Those who perform these tasks effectively receive handsome rewards, while those who do not quickly find themselves in a new line of business.

Successful companies continuously analyze market demands, technology performance, and resources to optimize their product lines. Competitors quickly integrate the key product features of industry leaders into their own designs based on consumer preferences and sales forecasts. Short- and long-term investments in R&D, production facilities, and support services undergo extensive performance analyses for financial (e.g., return on investment), technical (e.g., performance benchmarks), and business (e.g., market share) aspects. Businesses invest in data to regularly update and fine tune analytical models to support strategic and tactical decisions to maximize revenue, profits, and market share. They rigorously identify and prioritize market demands to exploit these opportunities with an optimal balance of portfolio solutions.

Time-to-market represents one of the most powerful drivers in commercial product development. Some companies seek to achieve “first mover advantage” by introducing a new product into the marketplace. Others then offer products or services with additional features, better performance, or a lower price point to gain market share. The more time that companies waste on perfecting “the next big thing” the more time competitors have to sell their products. Rarely are the best products on the market a business’s first version. Instead, an iterative series of competing models usually generates the strongest, innovative products, from the current year’s model hybrid car to the latest smartphone. While commercial enterprises operate in a different environment, the DoD can adopt many valuable private sector practices to structure and execute acquisition portfolios.
What Is a Portfolio?

A DoD portfolio would comprise a collection of programs, projects, increments, and related development efforts designed to achieve a set of strategic outcomes. A portfolio could expand on the system-of-systems model or span a program executive officer’s (PEO) full suite of programs. Many DoD headquarters organizations use portfolio management from a functional oversight perspective, rather than on designing integrated solutions. This portfolio vision is a more tactical approach to structure acquisition elements above a program by those closest to the program execution.

To avoid the common DoD pitfalls of complexity and bureaucracy, portfolios should encompass a small group of related programs, such as those within a PEO’s portfolio. For example, IT portfolios could manage a suite of applications and services that run on a common infrastructure platform, while aircraft portfolios could leverage a common airframe (e.g., C-130) with different payloads for each mission profile. Portfolios could also leverage common subsystems across programs, to include engines, sensors, communications suites, or avionics software (e.g., Special Operations helicopters). The DoD may find it easier to begin with portfolios of programs that are easily divisible, such as IT systems, rather than with large programs developing new bombers, ships, or space systems. Over time, if successful, the DoD could expand and scale these portfolios.

Overview of the Portfolio Acquisition Model

Just as industry has succeeded by applying a portfolio model around product lines, the DoD could achieve similar success by structuring and managing acquisition via portfolios. This would require decomposing large systems into multiple smaller programs, projects, or increments. These portfolios would group related capabilities across programs and commercial off-the-shelf (COTS) products and services, thereby elevating the time-consuming acquisition processes to the portfolio level, reducing program workload, and allowing programs to deliver products faster (see Figure 2).

![Portfolio Acquisition Framework](image)

**Figure 2. Portfolio Acquisition Framework**

A portfolio structure can foster innovation to deliver affordable solutions that achieve mission outcomes. The DoD would construct programs and increments from federated inputs, priority requirements, mature technologies, and rigorous analysis focused on a mission area. This would include an active government and industry R&D community aligned to advancing technology solutions. Enterprise management via portfolio strategies,
roadmaps, and architectures would guide development of a suite of smaller capabilities. Dynamic allocation of budgets, personnel, and other resources would lead the DoD to invest in the highest priority efforts. Portfolios would extend beyond delivery of an initial capability to optimize operations and sustainment of the capability suite.

<table>
<thead>
<tr>
<th>Table 1. 12 Major Elements of Portfolio Acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>REQUIREMENTS ANALYSIS</td>
</tr>
<tr>
<td>R&amp;D</td>
</tr>
<tr>
<td>STRUCTURE</td>
</tr>
<tr>
<td>ARCHITECTURE STRATEGY ROADMAPS GOVERNANCE</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>BUDGETS WORKFORCE CONTRACTING</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>SUSTAINMENT</td>
</tr>
</tbody>
</table>

**Acquisition Elements**

This section presents details on the acquisition elements shown in Table 1, describing the program model (as-is) and offering a vision for a portfolio model (to-be).

**Requirements**

**Program Model**

Programs capture initial requirements in an Initial Capabilities Document (ICD) at the start of the acquisition life cycle to outline a broad capability gap. They then refine and solidify requirements in a Capability Development Document (CDD) that contains key performance parameters. The Joint Requirements Oversight Council (JROC) must approve the CDD before system development begins. MDAPs usually take an average of 24 months to complete CDDs that in essence lock down the program scope for the next 10 to 15 years of development and production (Sullivan, 2015). During this time frame, change occurs constantly across operations, threats, priorities, budgets, technologies, and related systems, but the requirements remain fixed.

Operational sponsors often inflate the scope of a CDD by including all known requirements, as potential subsequent increment or program would follow many years later. This compounds risk by expanding the program scope, the number of critical technologies to mature, and variances in estimates, creating longer timelines to achieve initial operational capability.

**Portfolio Model**

Given the rapid pace of technology change, the DoD can no longer afford to lock in requirements for a decade or more. Instead of attempting to predict long-term operational and technical needs prior to defining short-term operational capabilities, programs must
focus on incremental advances. Managing via a broader set of portfolio requirements would enable greater system interoperability than a series of large, fixed CDDs for major weapon systems.

A dynamic and agile requirements model with users at its center would serve as the foundation for effectively scoping programs in a portfolio model (Figure 3). ICDs would cover a broad mission or capability area and align with the scope of a portfolio rather than that of a single program (Winnefield, 2015). They would be broad documents central to ensuring that the operational, acquisition, and intelligence communities align around common outcomes, priorities, and expectations. In coordination with operational commands, operational sponsors could manage capstone requirements via portfolio ICDs as living documents. This would include annual updates to reflect their current concept of operations, strategic guidance, priorities, threats, capability gaps, and desired effects.

A database or requirements management software would capture the next level of portfolio requirements, which many products or services in the portfolio could ultimately satisfy. A requirements board and team of operational sponsors would manage the dynamic requirements list, reprioritizing it on the basis of operational priorities, threats, and desired effects. The acquisition community and potentially industry would translate the items on the list into engineering requirements while exploring notional technologies and solutions for each. Portfolios would reprioritize and revisit the requirements regularly to ensure increased fidelity. The Defense Intelligence Agency could continue to supply inputs on mission and system threats as well as adversaries’ current and planned capabilities to help shape and prioritize requirements.

Programs and increments would have a smaller scope than today’s systems. The smaller the scope, the easier it would be to analyze, plan, estimate, design, develop, test, and produce capabilities with reduced technical and programmatic risks. Portfolios would scope the next program or increment on the basis of the highest priority requirements and the availability of mature technologies and affordable solutions. Delivering capabilities to users faster would reduce risk while responding more rapidly to changes in operations, technologies, and budgets. For example, portfolios would seek to deliver weapon system capabilities in five to 10 years rather than the 15 to 20 years common today and IT capabilities in less than 18 months rather than five to eight years.

To do so, portfolios would leverage the Joint Staff’s IT Box model, allowing more speed and agility in software system requirements (Winnefield, 2015). The IT Box model delegates requirements oversight and validation of documents following an ICD to a flag-
level organization rather than the JROC. Portfolios would streamline and tailor successor
documents according to the oversight authority and program needs.

**Analysis**

**Program Model**

During the Materiel Solutions Analysis (MSA) phase, programs conduct an AoA to
compare the operational effectiveness, suitability, and life-cycle costs of potential
alternatives. The Cost Analysis and Program Evaluation (CAPE) director provides guidance
for major programs and approves the final analysis. AoAs are led by the operational sponsor
with support from the acquisition community. The analyses often reveal a bias toward
alternatives that look and feel like the legacy system the new program will replace, but with
more modern technologies and improved performance.

Contrary to the perception that acquisition executives stress due diligence in this up-
front analysis, programs often experience pressure to complete the analysis so that they can
advance to the next acquisition phase in pursuit of the preferred alternative. Once a program
achieves Milestone A approval, it rarely revisits the AoA to validate constraints and factors
and ensure that the program is still pursuing the best solution. Programs refine their cost
estimates in each phase, with the life-cycle costs determined on the basis of tradeoff
decisions made early on.

**Portfolio Model**

Portfolio structures would enable robust, integrated, continual analysis to optimize
cost, risk, performance, and mission impact. Portfolio AoAs would be robust, continual
processes designed to optimize the performance and/or efficiency of a suite of programs
over their life cycle. Analysts would regularly assess the portfolio capabilities (fielded, in
development, and planned) to maximize mission impact and minimize portfolio life-cycle
costs. In-depth knowledge of technical baselines tightly aligned to cost models would drive
affordability and trade-space analysis at the program and portfolio levels. Portfolio-level
modeling and simulation (M&S) and experimentation would optimize system performance,
operational effectiveness, and suitability. Threat assessments would track adversaries’
military capabilities and the risk they pose to U.S. personnel, systems, and national
interests. These analyses would continuously monitor and evaluate a variety of
technologies, systems, services, and nonmaterial considerations such as doctrine, training,
or procedures. Technology advances would drive requirements changes and the resulting
system capabilities supported by a flexible contract and budget structure. Analyses of
programs in development would consider their acquisition performance and operational
priorities to ensure the programs continue to represent worthwhile investments. Data would
drive the design and adaptation of portfolio capabilities. Divestment analyses would assess
if and when to terminate a program and what alternative approaches to consider as a way
forward.

**Research and Development (R&D)**

**Program Model**

Programs in the Technology Maturity and Risk Reduction (TMRR) phase focus on
prototyping and maturing the technology to a point where the program can begin
development in the Engineering and Manufacturing Development (EMD) phase. Most
programs today develop the full scope of capabilities to meet all the approved requirements,
and the resulting systems can take a decade or longer to field. Individual programs are
responsible for maturing all critical technology elements and demonstrating them in a
relevant operational environment.
Program offices face pressure to transition to EMD as soon as possible so that they can deliver capabilities before the requirements and technologies become completely outdated. The Government Accountability Office (GAO) regularly criticizes the DoD for allowing far too many MDAPs to advance into EMD with immature technologies that create cost, schedule, and technical risks (Sullivan, 2014).

During the TMRR phase, many interested companies may contribute technology research and competing preliminary designs. Once a program reaches Milestone B, most R&D stops and a single prime contractor develops and produces the system.

**Portfolio Model**

A portfolio R&D environment would enable mission-focused research and rapid exploitation, both critical to maintaining technological superiority over adversaries. Establishing a long-term R&D environment for a portfolio would allow an active community to contribute to advancing innovative capabilities. Each portfolio could include government labs, federally funded research and development centers (FFRDCs), universities, DoD University Aligned Research Centers (UARCs), and diverse industry players in a collaborative environment. Portfolios could include pools of industry players large and small, traditional defense contractors, and innovative new entrants (see Figure 4). An open innovation culture would pursue ideas across contractors, partners, users, and even adversaries to shape R&D goals (Kelley, 2010). Both government and industry could contribute R&D funding to portfolio solutions and share intellectual property when appropriate. They would also make long-term investments in M&S, experimentation, and rapid testing capabilities. Portfolio leaders would provide their priorities for research and feedback to shape investments and determine which technologies to integrate into the next program. R&D organizations focused on technology maturity would reduce program risk and improve delivery speed.

![Portfolio R&D Environment](image.png)

**Figure 4. Portfolio R&D Environment**

As the DoD would increasingly rely on commercial technologies rather than military-unique developments, portfolios would make long-term investments in assessing current and emerging technologies. A portfolio R&D group knowledgeable about technology solutions would intelligently shape operations, requirements, and designs. This group would demonstrate capabilities, prototype emerging technologies, and compete in challenges to achieve performance goals. Robust M&S capabilities and experimentation would evolve, drawing on the latest technologies and threat assessments. Given the current era of
exponential technological growth, rapid and inexpensive testing would be critical for the portfolio.

**Structure**

**Program Model**

Huge, monolithic MDAPs develop all CDD requirements in a single, big-bang approach. MDAPs take 10 to 15 years from Milestone A to initial operational capability, with many of the largest systems taking even longer. Programs enter EMD with immature technologies, which leads to design instability, technical challenges, and significant cost and schedule overruns. Lengthy timelines between deliveries drive operational sponsors to add requirements to the scope of each increment, thereby compounding risks and increasing cost and schedule delays. For example, the F/A-22 took 22 years to become operational, with a 71% quantity decrease and 62% cost increase against initial plans. The Air Force could have delivered more capability sooner via a three-block incremental approach (Walker, 2013). The block upgrade model for B-52, F-15, and F-16 proved successful over decades, yet with its big-bang structure the F-35 program is struggling with costly retrofits.

In the year 2054, the entire defense budget will buy just one tactical aircraft … which will have to be shared by the Navy and the Air Force 6 months each year, with the Marine Corps borrowing it on the extra day during leap years. —Augustine’s Law XVI (Augustine, 2015)

**Portfolio Model**

Given competing missions, priorities, budgets, authorities, and many other factors, designing any element across platforms has historically added risk across programs, particularly joint programs. One of the biggest benefits of a portfolio structure would be the ability to design common platforms, subsystems, and services across programs. Stakeholders could shape these common elements to optimize portfolio performance, efficiencies, and mission impact.

Portfolios would structure developments to deliver a continual set of capability releases via small programs or increments. Smaller programs carry lower risk because of their well-understood scope, simpler design, more accurate cost/schedule estimates, and rapid delivery of capabilities. Speed reduces exposure to change and aligns requirements and capabilities delivered.

As illustrated in Figure 5, portfolios would scope each program or increment by leveraging the highest priority portfolio requirements and mature technologies from the portfolio R&D environment. This would help programs to deliver capabilities within five years for weapon systems and 18 months for IT systems, with estimated costs falling within the allocated budget.

![Figure 5. Bounding the Program Scope](image.png)
For example, instead of designing C4ISR aircraft independently, the DoD could examine the viability of a common aircraft platform with a modular design to allow for a diverse set of payloads. Common vehicles, communication suites, sensors, or ground stations would improve interoperability and cost efficiencies. Common services from IT infrastructure networks to system sustainment could improve mission impact and lead to cost savings.

Architectures

Program Model

Programs are designed individually and focus primarily on subsystem interfaces and performance. Each program develops a series of DoD Architecture Framework (DODAF) products to capture the capability, operational, services, and systems viewpoints (Winnefield, 2015). While these architecture products help programs to understand the bigger picture, designs remain program centric. A diverse set of defense industry contractors often integrates proprietary design elements, which creates risks to interoperability and system evolution. The maturity of architectures varies widely across the DoD, with few areas of a strong enterprise architecture driving program designs and interfaces. Programs have collaborated to jointly develop common subsystems, but often encountered considerable risk due to competing designs, distributed budgets, and cross-organizational dynamics. Many interfaces between systems are costly point-to-point designs difficult to evolve in a dynamic environment.

Portfolio Model

Establishing a portfolio for a mission area would provide a structure to develop and mature an effective enterprise architecture. Collaboratively developed and proven standards, interfaces, and processes would guide each program’s development. This strategic design approach would enable optimization in production, operations, and sustainment. A central portfolio authority for an enterprise architecture would ensure that new program designs leverage the architecture from the outset. Portfolios could more effectively design the modular open systems strongly advocated by Congress, the GAO, and the DoD’s Better Buying Power initiative (Kendall, 2014). Portfolio systems engineers would develop notional designs for each acquisition program using mature technologies from the portfolio’s development environment to address the top capability gaps identified in the relevant ICD. Robust portfolio enterprise architectures and collaboratively developed notional designs would outline how each capability fits within the portfolio suite. Portfolios would resist over-engineering complex architectures by driving simplicity and maximizing use of commercial technologies.

Strategies

Program Model

Major acquisition programs develop dozens of documents to support major milestone decisions. On average, programs take over two years to complete milestone documents, expending an average of 5,600 staff days (Sullivan, 2015). These documents force the program office to explore effective strategies for the next acquisition phase, yet the sheer quantity and complexity become overwhelming. As conditions change during the acquisition phase, programs rarely update strategy documents and resubmit them for approval. In short, program strategies are shortsighted and often do not reflect current approaches. Lengthy program strategies simply gather dust in file cabinets. “Working without a plan may seem scary, but blindly following a plan that has no relationship to reality is even scarier” (Fried & Hansson, 2010).
After awarding the contract, agencies are often locked into a single vendor for the program life. This eliminates competition—the single best method to control costs and improve performance.

**Portfolio Model**

Portfolio strategies would provide a long-term vision of how to deliver an integrated suite of capabilities most effectively and efficiently. The vision would include a clear set of portfolio goals, outcomes, risks, and performance measures. Unifying around an inspiring vision or challenge would provide clarity on investment decisions and rally a diverse community to develop innovative solutions. Portfolios should embrace LinkedIn’s CEO Reid Hoffman’s two rules for strategy decisions: speed and simplicity (Casnocha, 2015).

Consistent, repeatable processes across programs would foster a dynamic workforce, accelerate program execution, and allow for tailoring as necessary. Portfolio documentation would serve as the foundation for each program, thus reducing the amount of program-unique content to develop and coordinate. Common portfolio strategies and practices would ensure that each program leverages best practices and provide new programs an established framework on which to build.

Portfolio strategies would take industry considerations into account to optimize production lines across systems and foster an active, competitive environment. Integrating OSD/AT&L’s Sector-by-Sector, Tier-by-Tier (S2T2) industrial base analysis into program strategies would support a vibrant supply chain and affordable, stable development and production rates (Manufacturing and Industrial Base Policy [MIBP], n.d.). Strategies would explore innovative approaches to nurture an active industry community in R&D and in program development/production, and would consider sponsoring competitions to address critical risks or opportunities. Strategies could encompass more dual awards, split buys, and parallel developments to keep participants in an active contractor base leapfrogging each other with evolutionary upgrades or new, revolutionary solutions.

**Contracting**

**Program Model**

Contracting today involves a set of lengthy processes, with source selections that often take a year or more to complete. The contractor or contractor team selected to design and develop a new system often gains monopolistic power over the government for a majority of a program’s life span. As the DoD has moved toward acquiring larger and fewer major systems, this has changed the dynamics of the defense industry. Instead of creating a steady pipeline of potential work through periodic competition for new work, many of these large contracts become all-or-nothing, make-or-break outcomes that shape a major market segment for a decade or longer.

**Portfolio Model**

Portfolio contracting would focus on developing active, long-term partnerships with many companies rather than only a few. The goal would be to build a vibrant community of large and small companies actively contributing to R&D, architectures, designs, development, production, and sustainment of portfolio capabilities.

Competition remains the best way to drive down costs and increase innovation in defense programs. Therefore, a portfolio strategy should actively foster continuous competition over a program’s life cycle via broad industry participation. Decomposing large systems into a smaller set of programs would increase opportunities for industry, especially small businesses, to compete for DoD work. A potential portfolio contract strategy could use
multiple-award, Indefinite Delivery/Indefinite Quantity (IDIQ) contracts to establish targeted pools of large and small businesses with key technological and domain expertise.

The DoD could streamline contract timelines by establishing portfolio contracts with standardized business practices and pre-competed contract vehicles to enable rapid generation of task orders for programs and increments. Standardized business practices would include pricing, terms and conditions, templates, and selection criteria. Portfolios could maintain continuous competition by restricting the size of the contract vehicles with on and off ramps to refresh the vendor pools. Past performance on task orders within the portfolio would represent a valuable selection criterion for future work, as it would reward superior performance by contractors.

A portfolio approach should incentivize innovative companies to pursue defense work. New entrants, more than the major defense companies, offer the greatest promise for designing and integrating technologies in new ways to achieve a military advantage. The DoD has a variety of contracting programs to reach companies willing to offer new technologies, collaborative research, and experimentation. Broad Area Announcements (BAAs) foster competition to advance state of the art research and prototypes. Small Business Innovative Research (SBIR) and Small Business Technology Transfer (STTR) programs fund cooperative R&D projects with small businesses and universities (DCAA, n.d.). Portfolios could provide these small companies with an environment to prototype and demonstrate a focused set of capabilities tightly aligned with an operational mission. Promising small businesses could partner with established defense companies to navigate the DoD's regulatory gauntlet to develop and produce a new system.

Roadmaps

Program Model
Each program must develop and maintain a strategic schedule and detailed integrated master schedule (IMS). The quality of program schedules often increases in the lead-up to major milestones, while dropping off during acquisition phases. Detailed IMSs should integrate government and contractor activities, yet are often managed as contractor deliverables. Some operations, acquisition, and budget headquarters may have roadmaps or enterprise view of program schedules, yet the underlying data often lacks sufficient fidelity or currency.

Portfolio Model
A portfolio roadmap such as the one shown in Figure 6 would serve as a central, long-range planning tool for operations, acquisitions, and budget domains and include the following:

- Schedules of all legacy systems and planned programs/capabilities
- Quantities of operational systems and new production planned
- Identification of gaps, overlaps, and migrations from legacy to modern systems
- Current and projected performance levels for systems or mission areas
- Identification of legacy system risks due to technical factors, sources, or O&S costs
Portfolio roadmaps would provide operational, acquisition, and budget leaders and stakeholders with an integrated plan. They would support collaboration across these domains on status, risks, and plans, and foster discussions on priorities. Identifying risks or gaps would support decisions on accelerating new systems, delaying retirement of legacy systems, or implementing interim fixes. Aligning roadmaps with portfolio cost estimates and budgets would enable portfolios to optimize investments, ideally supported by analytical tools and methods. Many leading schedule software products already enable linking of program schedules. A portfolio schedule framework that integrates program dependencies would show the impacts of program schedule slips and support scenario planning.

**Governance**

**Program Model**

Governance presents one of the biggest challenges to effective portfolio management. Different stakeholder organizations across domains and levels have a competing set of priorities, incentives, cultures, and constraints. PEOs oversee the execution of the individual programs in their portfolio, but dedicate little time and resources to cross-program integration and optimization. The larger the portfolio, the harder it is to manage. Each layer of oversight across requirements, acquisition, and budget communities and functions groups programs differently, with little alignment around common portfolios. In some instances up to 56 organizations at eight levels reviewed program milestone documentation (Sullivan, 2015). With no two portfolios the same, it is difficult to reach consensus across communities on program priorities and budgets. The DoD incentivizes program managers to ensure that their program delivers the required performance within cost and on schedule. External dependencies are seen as risks. Therefore, many PEOs believe that the best way to minimize risk consists of scoping each program to include its own infrastructure as well as all subsystems and support equipment. Each program then progresses through the acquisition life cycle on an individual schedule and meets with its Milestone Decision Authority only at major milestones.

**Portfolio Model**

Portfolios would govern through collaborative, strategic partnerships with five key elements:

- Shared responsibilities of operational, acquisition, budget, and sustainment executives
- Portfolio alignment to ensure stakeholders represent the same mix of programs
• Decision authorities delegated to the appropriate level to enable timely decisions
• Central knowledge repository to provide stakeholder transparency and leadership insight
• Incentives aligned to ensure all organizations are working to common outcomes

Carefully limiting portfolio scope would ensure a manageable governance level. Program managers should be empowered to make decisions about technologies and subsystems (Berteau, 2014). Regular discussions among a diverse stakeholder group on priorities, status, risks, resources, and opportunities would ensure that the pipeline of programs supports the desired portfolio outcomes. Partnerships between operational commands and acquisition portfolios would foster collaboration on operational details and on which technologies/capabilities can be rapidly tailored for their missions. The partners would have wide latitude to shape the program scope and features.

Portfolio charters would clearly define authorities, roles, and responsibilities. Online repositories would capture and share portfolio knowledge to provide real-time insight and a common understanding. Embracing a servant leadership mindset would foster program support, integration, and innovation. Robust portfolio analytics would enable the data-driven decisions essential in these complex environments.

Governance would balance gate-check reviews (e.g., milestones) with time-phased portfolio and program reviews. Establishing a battle rhythm to discuss program status, issues, and ways ahead would minimize the burdens imposed by major milestone reviews. Portfolio strategy documents would reduce the burden on programs by requiring only constrained annexes that contain program-unique information. Reviews would still take place to ensure programs have a sound enough strategy and mitigated risks to warrant entry into and funding for the next acquisition phase. With delegated authorities, common processes, and regular insight, programs would minimize the documentation and reviews required to make informed decisions.

**Budgets**

**Program Model**

Historically, the DoD programs with the largest budgets have been the most likely to overrun costs and fail to deliver capabilities, while those with the smallest budgets were most likely to underrun cost and exceed performance expectations (Ward, 2014). Most acquisition programs today are funded via budget accounts called program elements (PEs), which are outlined in the president’s budget to Congress and included in the annual appropriations bills. Funding for each program is closely monitored by Congress, the DoD comptroller, each service and agency, and the program managers. PEs often fail to provide Congress with consistent, complete, and clear information (Sullivan, 2007). The lengthy DoD Planning, Programming, Budgeting, and Execution (PPBE) process requires budget requests and approvals years before programs are executed, with frequent adjustments made each year. The biggest challenge posed by the current budget constraints involves responsiveness to changes in operations, threats, opportunities, program performance, and priorities. Transfers of funds between PEs are limited to 10% of the budget for the current execution year, with congressional approval needed for larger transfers.

**Portfolio Model**

Some PEs today include multiple programs, with each broken out at a subaccount level called a budget program activity code (BPAC). Transferring funds between BPACs
requires lower approval thresholds than transfers between entire PEs. Thus, allocating a portfolio budget at the PE level with programs at the BPAC level would offer funding flexibility and agility, while also providing sufficient transparency to oversight officials.

This funding approach would increase the effective use of constrained resources and would direct funds toward the highest-priority capabilities with the greatest enterprise impact. Pentagon executives would focus on strategic budget allocations at the portfolio level. Portfolio stakeholders would allocate program funding following key milestone reviews. Portfolio managers would then establish funding lines for technology development, enterprise platforms, and personnel for enterprise efficiencies. Fortunately, such a change would not require a wholesale restructuring of the PPBE process but would simply call for shaping a few PEs for an initial set of portfolios.

Workforce

Program Model

Program office staff are often assigned to a single effort for an extended period of time, limiting their exposure to and experience with other programs or DoD-wide procedures and often leading to atrophy of their skills. Military personnel rotate every three to four years, with program management turnover frequently highlighted as a systemic program risk. While stability of key leadership positions can be beneficial, an inflexible staffing model that ties staff to a program for a decade is grossly inefficient and ineffective.

Hundreds of acquisition programs go through roughly the same major acquisition processes, yet often reinvent the wheel each time rather than tailoring a common approach to program specifics. As a result, a program planning for a major event—for example, a Preliminary Design Review (PDR)—may have few staff with recent PDR experience, and most staff may need to relearn some of the key elements to prepare for and execute the PDR.

Portfolio Model

Programs and acquisition workforces would perform more efficiently and effectively in a portfolio matrix organization. In an era of budget and workforce challenges, a dynamic staffing model would yield cost efficiencies, a strengthened workforce, and improved program outcomes.

Each program would have a balance of long-term staff with deep historical program knowledge and technical and process subject matter experts (SMEs) dynamically assigned throughout the program’s life cycle. In a portfolio structure, individual programs or increments would have shorter durations, which itself would reduce the skill decay that can result from lengthy program assignments.

In a portfolio matrix model, a percentage of the workforce could serve as process or technical experts who augment program office staff via short-term assignments. Process experts, for example engineers who specialize in system design, could advise program offices in the preparation and execution of PDRs and Critical Design Reviews and their associated design drawings. Schedule experts could assist in development and implementation of integrated master schedules to effectively manage the program and its dependencies on external efforts. Market research or commercial technology experts could ensure programs have a sound understanding of market offerings and technology solutions to shape the program scope and strategies.

Technical experts, by contrast, would offer deep insight in particular technical domains (e.g., avionics, sensors, stealth, or cyber). As programs progress through the
acquisition life cycle, these SMEs would phase in and out of the program office as conditions warrant. Using expertise only when required, instead of committing personnel to long-term assignments while demand for specialized skills ebbs and flows, would provide an optimal staffing model. SMEs could support multiple programs at the same time, thus establishing repeatable processes and horizontal integration across the portfolio.

Process and technology SMEs would focus on mastering their niche areas by collaborating with other SMEs across the DoD. Process SMEs would develop and maintain guides, templates, and repeatable processes for easy program adoption. Technology SMEs would research and collaborate with labs, FFRDCs, and industry in a focused technology domain to support program designs and innovative solutions. As staff members progress through their careers, they could transition between program and process focused roles.

**Sustainment**

**Program Model**

Government depots and prime contractors sustain the DoD’s weapon systems following a variety of operational models. Related major programs in a similar mission area are often sustained at diverse locations across the country, leading to massive inefficiencies in facilities, personnel, and support equipment.

**Portfolio Model**

Portfolio enterprise architectures and designs would enable strategic sustainment strategies to leverage common subsystems, parts, and support services. Portfolio sustainment strategies would leverage economies of scale via strategic investments and operations. Designing a holistic approach to sustaining portfolio capabilities would enable government and industry to make smarter long-term capital investments for production and sustainment. Subdividing monolithic systems into capability suites would create a smaller, steady pipeline of new systems to sustain. An enterprise analysis of costs, benefits, and risks could support a balanced portfolio of leasing versus buying solutions. Portfolios could establish public–private partnerships across programs, considering resources, demand, and expertise. Portfolio-level sustainment performance metrics and measures could incentivize industry to move from system-specific measures toward integrated mission-area capability rates.

**Summary**

Acquisition programs today are burdened by the complexity of the acquisition environment, the difficulty of maturing critical technologies, and the inability of the acquisition system to respond to changing operations, technologies, and budgets. Budgetary, workforce, and regulatory constraints further compound program risk. In a complex, integrated environment, defense acquisitions can no longer rely on a structure based on individual systems. Embracing a capability-focused portfolio structure modeled on the commercial sector offers many solutions to the DoD’s top challenges.

The principles of simplicity, commonality, and agility should guide all acquisition portfolios. By adopting the commercial product-line approach, the DoD could address long-standing acquisition issues associated with speed, resilience, and interoperability. Elevating the time-consuming acquisition processes to the portfolio level would reduce program workload, allowing each program to deliver products faster. Managing requirements, budgets, and staffs at the portfolio level would enable dynamic allocation to high-priority programs. Portfolio strategies, roadmaps, and architectures would guide program development.
In a portfolio structure, an active government and industry community would collaboratively develop technologies and designs and employ continuous competition to develop and produce the individual systems. Portfolios would design and optimize acquisition processes to deliver a suite of smaller programs rapidly, ensuring that warfighters regularly receive integrated, incremental capabilities with the latest technologies designed to achieve operational missions.

Advancing a portfolio model will require the DoD to address various strategic challenges. Congress maintains strict control over program budgets and location of depots to sustain systems. Reaching agreement between the DoD and Congress on the proper balance of insight, authorities, and accountabilities will take time. Each functional area (e.g., requirements, systems engineering, testing) would require tailored processes and possibly new policies to enable portfolio strategies, and the DoD would need to identify which suite of programs would comprise the initial portfolios. Finally, the culture of the acquisition workforce would have to shift to support a new portfolio acquisition model. With forward-thinking acquisition leaders in place across the Pentagon and Capitol Hill, the DoD has a prime opportunity to pursue a portfolio acquisition model that can achieve transformational solutions.

References


Sullivan, M. J. (2007). Defense acquisition: DoD’s research and development budget requests to Congress do not provide consistent, complete, and clear information (GAO-


**Panel 18. Promoting Secure and Timely Development and Sustainment in Weapons System Procurement**

<table>
<thead>
<tr>
<th>Thursday, May 14, 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>3:30 p.m. – 5:00 p.m.</td>
</tr>
<tr>
<td><strong>Chair:</strong> Reuben Pitts, President, Lyceum Consulting, LLC</td>
</tr>
</tbody>
</table>

**Towards Rapid Recertification Using Formal Analysis**
Daniel Smullen, Carnegie Mellon University
Travis Breaux, Carnegie Mellon University

**Acquiring Secure Systems Through Information Economics**
Chad Dacus, Air Force Research Institute
Panayotis Yannakogeorgos, Air Force Research Institute

**Practical Application of Principal Agent & Auction Theory to Improving Acquisition Outcomes: A Contracting Guide**
William Griffin, Justin Blott, Nicholas Boardman, Andrew Cady, Jake Elliott, Nick Mastronardi and Parker Quinn, USAF, Air Force Academy

**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.
Towards Rapid Recertification Using Formal Analysis

Daniel Smullen—is a research assistant enrolled in the software engineering PhD program at Carnegie Mellon University. His research interests include privacy, security, software architecture, and regulatory compliance. [dsmullen@cs.cmu.edu]

Travis Breaux—is an assistant professor of computer science in the Institute for Software Research at Carnegie Mellon University (CMU). His research program searches for new methods and tools for developing correct software specifications and ensuring that software systems conform to those specifications in a transparent, reliable, and trustworthy manner. This includes compliance with privacy and security regulations, standards, and policies. Dr. Breaux is the director of the CMU Requirements Engineering Lab, is a co-founder of the Requirements Engineering and Law Workshop, and has several publications in ACM- and IEEE-sponsored journals and conference proceedings. [breaux@cs.cmu.edu]

Abstract

Department of Defense (DoD) acquisition requires IT to undergo the DoD information assurance certification and accreditation process (DIACAP), which makes architecture-dependent assumptions. Emerging IT architectures, such as mobile and cloud-based platforms, invalidate these assumptions and prevent the DoD from acquiring commercial technologies that are readily available to adversaries. To address this problem, we extended our initial automation framework, wherein an application profile is expressed in a formal language and scaled with evolving architectural assumptions. These profiles will help ensure that information assurance requirements are commensurate with risk and scalable based on an application’s changing external dependencies. Information assurance risk levels must account for changing environmental and IA parameters (confidentiality, integrity, and availability) that result from dynamic recombination of applications during runtime. Our proposed language aims to address dynamically composable, multi-party systems that preserve security properties. Software developers and certification authorities can use these profiles expressed in first-order logic with an inference engine to advance the DIACAP and re-check compliance as IT systems evolve over time.

Introduction

Ensuring confidentiality in information systems is of paramount importance to mitigate the likelihood of data spills. But as systems change and requirements evolve, it quickly becomes unclear how these changes may affect the security of protected information with regard to secure enclaves. The DoD is increasingly reliant on software in all operational contexts, and agility in terms of the ability to certify and deploy new and improved software technologies necessitates greater agility in certification processes. Software recertification processes require significant expenditure in order to provide evidence of information assurance (IA) policy conformance. The costs of both sourcing and developing software are compounded by the need to maintain these certifications, especially when changes occur. Current processes are manual, and cannot scale as complexity increases. Savings and scalability can be achieved by employing formal analysis to assist humans and manage risk. This increases trust and reduces validation time, useful in complex systems and where high-confidentiality data may be deployed in low-confidentiality enclaves (or in other high risk scenarios).

We believe rapid recertification can be enhanced by documenting assumptions in tool-supported frameworks where assumptions that continue to hold may be reused. One approach is to use lightweight formal analysis to model and validate security requirements. This analysis should focus recertification efforts on only those requirements that may conflict with IA policy. Furthermore, such conflicts should be resolvable through reconciliation...
strategies in which changes to the system can be checked against conformance to the IA policy, without requiring a complete review of the entire system.

In this paper, we present our methodology which principally focuses on modelling and validating specifications of data flow as the basis for evaluating IA policy. The method permits automated analysis of data flowing into and out of a system or component to detect conflicts in the data’s specified purpose, which illuminates potential areas of non-conformance and further expedites conflict resolution in regard to the recertification process. We also classify and present reconciliation strategies for resolving the different types of conflicts that may occur. We show that our method is scalable to permit analysis of large, complex, and evolving systems—whose specifications can involve numbers of policies no longer tractable by manual analysis alone.

Organizational granularity refers to the view at which a policy or design artifact is intended to represent its context within an organization. Our scalable, automated analysis can be repeated rapidly as specifications change with the aim to reduce recertification time at any point in the software lifecycle and at any level of organizational granularity. For example, design artifacts detailing collections of software components represent a very high degree of organizational granularity. Networks and departmental interconnections represent a medium degree of organizational granularity. Enterprise-level architectures and inter-organizational connections represent a very low degree of organizational granularity. From the software perspective, these conventions are often referred to by their "level," which comprises the same notion as organizational granularity but on an opposing scale: Software components and their implementations are referred to as low-level artifacts. Network diagrams, detailed design documentation, and inter-departmental processes are referred to as mid-level artifacts. Inter-organizational processes and enterprise architectures are referred to as high-level artifacts. Discussions which must be paired with their context for policy or design artifacts in this paper are referred to by their organizational granularity, for consistency.

Changes to information systems include the addition of new features or new behaviors, or the establishment of new connectors between existing system components and others. Thus, we can envision a range of scenarios that yield characteristically different recertification challenges. For example, designing new systems to replace legacy systems (early lifecycle), integrating new systems with existing systems (mid-lifecycle), reworking existing systems to perform new functions (mid- to late-lifecycle), and during perfective maintenance tasks (late-lifecycle). During these lifecycle stages, software systems are increasing in size due to the nature of these changes, and this increases the time and cost for validation and recertification.

Unfortunately, the naïve approach of dividing up the software system and parallelizing recertification tasks is a separate and more costly challenge that does not gain traction over the problem. Decomposing the system in ways that violate architectural and source code artifact boundaries, such as software interfaces, can increase accidental complexity of certification tasks (Brooks, 1995) when systemic quality attributes, such as ensuring confidentiality, cut across these boundaries. For this reason, our approach employs the notion of a context that may cover a specific application, called an application profile, or an entire secure enclave. Within this context, we can reduce data flow analysis from across enclaves to conform to the same expression and reasoning needed to detect policy conflicts within an enclave. This level of reasoning would correspond to a high or medium level of organizational granularity. In Figure 1, we present two secure enclaves, “A” and “B,” with a support service connected to enclave “A” and a handheld application connected to enclave “B.” The links that connect these two enclaves are assumed to be
secure as they are within the same operational environment. However, bringing in support services and mobile devices from the outside context into these secure environments represents a change in functionality, which may be a recertification trigger. These recertification triggers are discussed later in our Evaluation section.

Figure 1. Example of DoD Distributed IT System Illustrating the System Decomposition Boundary Between Secure Enclave and External Areas

In this simplified context, we focus narrowly on data actions that access and make available information and postpone addressing questions about which specific security mitigation is needed to address a specific vulnerability. For example, answers to questions about when to use encryption correspond to sensitivity of data and under what contexts that data is made available, which we account for. However, the question about what level of encryption to use is not a central focus of our method, and is well documented in IA policy based on data sensitivity. We now discuss the technical background to our approach.

Technical Background

The Bell-LaPadula model simplifies the characterization of information flow from low confidentiality to high confidentiality, but not vice versa (Bell, 2005). This model has long been the traditional view of enforcing multi-level security policies in government and military applications. In our method, we characterize the purpose for which information is used as the security level and then allow policy authors to express compositions of security levels through containment and disjointness, for example, a security level may contain or be disjoint with another level. This formalism extends our prior work on semantic parameterization (Breaux, Anton, & Doyle, 2008) for expressing actions on data in Description Logic (DL) as a composition of actors, objects, and purposes, and for transactions involving data, the source and target of the transaction. More recently, we developed a human-readable SQL-like language for expressing these application profiles (Breaux, Hibshi, & Rao, 2014), which we refer to in this paper as the “Application Profile Language.” Application Profile Language syntax is parsed and compiled into the Web Ontology Language (OWL), which is suitable for computer processing by an automated DL reasoning tool (Bechhofer et al., 2004).
Subsumption is a syllogism in which one concept describes a more general class of another. This is commonly referred to as a superclass/subclass relationship. In DL, we can check whether one data action subsumes another action, which means every interpretation of the second action is contained in the set of interpretations of the first action. Subsumption allows us to detect conflicts, in particular, when one action is deemed permissible and the same or a subsumed action is deemed impermissible. The relationships between concepts expressed in a specification are mapped directly into the DL model. For example, through subsumption, high-confidentiality purposes for a data transmission may include “top secret” or “for mission-critical purposes,” or any number of other concepts which are desired to express with respect to crafting a data requirement.

The DL model is comprised of two parts:

1. an ontology in which key terms are defined, including information type categories and type compositions expressed using subsumption
2. a set of rules governing collection, usage, and transfer of data to third parties

For each data action, the purpose for the action and the party from whom that data is sourced is stated, and for transfer, the party who will receive the data is stated. Actions may be expressed as permitted or prohibited in the application profile. The rules expressing permissions or prohibitions in the profile represent the high-level specification of what actions an application is permitted or restricted from performing, whereas the implementation would entail mapping these actions to functions in code, such as database queries or radar telemetry-based analysis, and so forth.

Figure 2 is a process diagram which illustrates the recertification process steps using our tool. The highlighted area on the right of the figure spotlights the novel contribution of this paper, which is the conflict reconciliation strategies we present later. The recertification challenge that we focus on herein concerns how to compose systems of systems that collect, use, and transfer data across system boundaries, between secure and unsecure enclaves.
Approach and Running Example

We now present a running example that we use throughout the paper. The example draws from public accounts of real-world vessels and technologies that are currently undergoing sea trials—specifically, we chose scenarios regarding Zumwalt-class destroyers, which we generalize and refer to as “Zumwalt.” The running example is used to illustrate the high-level certification concepts that exist in real software systems aboard sophisticated platforms, such as those integrated into Zumwalt’s Total Ship Computing Environment Infrastructure. Certification concepts deal with software requirements, and are not reasoned about based on lines of code, given the complexity of software. Zumwalt’s electronic systems are comprised of approximately 6 million lines of code—far too much to reason about at the low level, pedantically analyzing each line to determine if it meets requirements or not. Software must be analyzed in terms of high-level concepts, and certification auditors must align analysis methods to meet the same high level of abstraction.

To help achieve alignment between the level of abstraction for our analysis and high-level abstraction in software, details unrelated to software requirements about the ship-class have been removed. Details about the software requirements themselves have been generalized in order to make the focus of our discussion centered on the essential transitive qualities of data. We are not concerned about the underlying functionality or supporting systems of compartments aboard; rather, we are focused on understanding and analyzing systemic behaviors that specify to consume or move data to a certain place, for a certain purpose, to achieve requirements. Rather than being concerned about the memory chips, logic boards, networks, and radiation-emitting or receiving hardware that comprise the air- and surface-borne vessel sensing capabilities aboard a ship, we can abstract this entire sensory capability as the radar system. We can reason about the broad concepts of radar data that may be produced by such a system. The application profiles which we use to formalize such concepts give us the capability of automated reasoning at this level. In Application Profile Language, the radar system would be written as an actor. The radar data would be expressed as a data concept. Requirements, expressed through policies inside application profiles, provide the necessary details for determining what is being done with this data in the context of expressed actors. The contents of an application profile can therefore be articulated in notional scenarios based on our running example that captures the necessary details to express an intended new system functionality.

Our running example is designed to be illustrative of high-level software concepts, and may not reflect the entirety of the underlying details of a vessel, the bits and bytes of data in their onboard computer systems, or a platform’s true capabilities. However, the example provides sufficient detail about where and what data is being used, so that we can reason about data’s transitive properties given the general characteristics of these systems. Without worrying about the lines of code that comprise these software systems, we can reason about how data is used, and how this data may need to be transferred outside of a vessel’s enclave to another party to fulfill requirements.

In our running example, Zumwalt’s key capabilities that we have abstracted correspond to some of the platform’s distinguishing features compared to other destroyers; these include Zumwalt’s advanced radar system (Tolley & Ball, 2014), abstracted as the radar system. Zumwalt’s sensor-netting capabilities that permit sharing of information with friendly platforms (O’Neill, 2007), including AEGIS technologies or the Total Ship Computing Environment Infrastructure, are abstracted as the information sharing mechanism. Together, these two abstracted systems permit enough expressivity in terms of the data concepts that exist in the real world so as to maintain realism with respect to DoD IA policies and software certification processes. We reason about requirements for these computerized systems...
using already established directives for security used in practice, such as DoD Directive 8500.01. Zumwalt may be considered to be in any theatre, with no specific mission, which permits our example to be illustratively modified and re-examined in different operational contexts. In any of these contexts, the underlying thought processes that are required to analyze decisions about adding new features, or modifying connections with outside information systems and processes, are the same. This focus on only information sharing concepts rather than technical implementation details also means that these scenarios could be extended to cover any vessel, or land-based system.

The Zumwalt-class’ software implementations have made extensive use of real-time operating systems with stratified and segmented “high” and “low” security networks as part of the Total Ship Computing Environment Infrastructure (Lynxworks, 2007). These technologies conflate directly with the Bell-LaPadula (Bell, 2005; Landwehr, 1981) view of IA and security, and is the same model for purpose which we show in our DL ontology. Attention on software aspects of the class’ radar system and information sharing capabilities is an identified concern in design reviews of the vessel (O’Rourke, 2012), emphasizing the need for reducing software certification costs and time, which contribute to the expenses associated with information sharing capabilities such as sensor netting. Certification becomes an even larger problem when considering information integration with multiple platforms or allied vessels in the demanding context of future battle spaces, which is the intent behind our selection of Zumwalt as the subject of our example.

Profiles, Conflicts, and Tracing

We now describe our approach to trace data flows in and through application profiles, and to detect conflicting requirements. When policies covering multiple organizational granularities are used in a single application profile’s policy, overlapping policies may result. In high and medium granularity overlaps, organizational policies may conflict with individual policies governing software components, and vice-versa, for example. In low and medium granularity overlaps, inter-organizational policies may conflict with departmental policies. High and low granularity overlaps are also possible, in which inter-organizational policies may conflict with those governing software components. In general, overlapping policies may lead to conflict, and these conflicts must be reconciled in order to prevent data spills. Later in this paper, we discuss identifying and reconciling specific types of conflicts that exist as a result of these overlaps in light of this risk.

Actions on Data

In our model, application profile rules may govern three primary actions which express the transitive nature of data as it moves through a system. These actions are collection, usage, and transfer of data. Each action concept has assigned roles that relate the action to actors, data, and an associated purpose. The collection action describes an act by a party to access, collect, obtain, receive, or acquire data from another party. The usage action describes an act by a party to use or consume data in any way. Transferring describes an act by a party to transmit, move, send, or relocate data to another party so that they may collect it. Pairs of rules that permit collections and transfers on the same datum are referred to as a data flow. When one of the rules in the pair is expressed with respect to a third party, this data flow can be traced to the third party.

For example, Zumwalt may need to transfer information about the presence of an enemy radar contact to a friendly vessel in the vicinity in order to initiate a combined engagement. In another scenario, a friendly vessel may need to transfer the same information to Zumwalt so that Zumwalt can engage it. In both scenarios, these actions performed on the radar contact data represent collection, transfer, and usage actions.
Collection actions generate radar data from both vessel’s radar systems, and permit one vessel to receive data from another. Transfer actions permit one vessel to share its data with another. Usage actions permit a vessel to engage the radar contact once the data is within its enclave.

Actions are further described by DL roles. The hasDatum role represents the action’s affected datum. The hasActor role indicates the source actor from which the data was collected. The hasPurpose role indicates the purpose for which an action is being performed. These purposes, in line with the original Bell-LaPadula model, are abstracted as high-confidentiality and low-confidentiality. These classes of purpose may be further subsumed by any other purpose in order to extend the ontology to fit the policy expressivity needs of a specific organizational granularity.

Finding Conflicts in Policies

We now describe how the rules governing actions on data may conflict, and how we can detect these conflicts automatically. In our DL ontology, a conflict is defined as an instance when an action is permitted by one rule, and prohibited by another. Rules governing the permission and prohibition of specified actions are described using the Application Profile Language, which is parsed by our language parser into Web Ontology Language (OWL). OWL can be analyzed automatically by freely available open source DL inference engines such as Pellet, Fact++, Racer, and HermiT. These DL reasoning engines have historical pedigrees in academic usage for formal analysis (Baader, Horrocks, & Sattler, 2005).

In the DL representation of an application profile, a rule must be determined by the reasoning engine as equivalent to both a right (also known as a permission) and a prohibition, in order to be found equivalent to a conflict. Both the right and the prohibition must act over the same datum and for the same purpose in order to be equivalent.

An equivalence relationship requires equivalence in both directions, as one might find in a mathematical expression (such as \(a + b = c\)). In some cases, the conflict can only be reasoned about with respect to one rule, because reasoning can only occur on “one side” of the equivalence operator. This is explained through subsumption. We cannot distinguish the rule on the other side of the equivalence as truly involved in a conflict; there may be further subsumed interpretations using subclasses which do not conflict.

Using our running example, we may imagine an application profile which governs how Zumwalt is permitted to share radar data that it has collected. In this case, radar data contains data about enemies, friendlies, and terrain. In DL, radar data subsumes data about enemy vessels, data about terrain, and data about friendly vessels. When radar data subsumes these more general concepts of data, radar data is the subclass, and data about enemies, terrain, and friendly vessels are the superclasses. The DL reasoning engine must reason about all of the superclasses of a datum when reasoning about the subclass, and different rules may exist that govern each of these data.

We instantiate this example with a profile that contains data definitions and a basic policy for Zumwalt’s radar and information sharing, seen in Figure 3, and conforming to the profile language syntax defined in Breaux et al. (2014). This figure also shows the abstraction of Zumwalt’s radar system and information sharing system as actors. On the top of the figure are the data definitions that correspond to the header section of the profile. On the bottom of the figure is the policy, which is comprised of five rules. Together, the header and the profile of a policy comprise a full application profile. Each data definition and rule seen in Figure 3 is annotated with the English language representation of the Application Profile Language syntax.
This profile has been seeded with a conflict that is derived from the intended system functionality that would permit Zumwalt to share collected radar data with friendly fleets for low confidentiality purposes. The intention of this policy is to provide information sharing on this vessel, but this conflicts with an overarching organizational policy that does not allow data about friendly vessels to be shared with other friendly vessels. We now describe the nature of this conflict and how it is detected by our approach.

Figure 3. Representing Zumwalt as Two Abstracted System Components (Actors), With an Associated Application Profile That Describes the Rules That Govern the Data Produced and Transferred by These Actors

If we write a policy that permits us to share all radar data, but prohibits sharing data about friendly vessels, we cannot share all radar data. This is a one-sided conflict. This conflict occurs because the relationship between these general data concepts and the specific radar data subclass is actually defining a composition rather than a subsumption; however, the semantics of the relationship are the same when expressed in this way. As a result, we may not be able to reason about the superclass of general data that we have defined here, but we can reason about the subclass of radar data, and its relationship with friendly data. Superclasses and subclasses are not equivalent, but we can infer enough about radar data to know that the policy would still generate a conflict—this reasoning about the subclass rather than the superclass is similar to reasoning about one side of an equation based on inferences about the other side (hence the name one-sided conflict).

One-sided conflicts are highlighted by the DL reasoning engine on one rule. This is convenient for our purposes because this single rule is the one that must be acted upon using the conflict reconciliation methods seen in the next section. The other conflicting rule, governing the superclass, cannot be reconciled using the approaches we have identified.
due to the nature of this one-sided conflict. We cannot say for certain whether we should perform the reconciliation action on the “other side” of the conflict since we cannot say that it conflicts in all possible subsumed interpretations. An example of such a conflict is seen in our running example.

We infer that collection/transfer data flow traces are functionally the same as collection/usage traces, in terms of the transitive properties of data. However, given that the scope of our analysis is to determine where data spills may occur due to information transmission to third parties, collection/usage data flow traces are not relevant. Collection/usage flows can only occur within the system bounds of a single party. A party may consist of more than one secure enclave. Collection/use traces may still generate conflicts, but these conflicts would indicate mismatch of an intended purpose within a secure enclave, and do not impose a serious risk of data spills unless there is a separate rule permitting the transfer of this datum elsewhere.

**Conflict Reconciliation**

When conflicts arise, we have identified two main strategies which work to match the generality of purpose. Matching the level in the class hierarchy of subsumed purposes reconciles conflicts in which a more specific high-confidentiality purpose is permitted, but more general low-confidentiality purposes are prohibited. This serves to mitigate the likelihood that data will be transferred outside of the secure enclave without explicit authorization for a specific high-confidentiality purpose. Matching purposes also serves to mitigate being mismatched with a third party that will consume (or retransfer) this data for a more general purpose, which would constitute a data spill. The reconciliation actions which we have identified, redaction and generalization, serve the same purpose as their namesakes in legacy document-oriented processes. Both actions serve to transform data in such a way that it is permissible to transfer it for low-confidentiality—and therefore more general—purposes.

**Redaction**

Redaction means to remove elements from a collection of data in order to limit the spread of information that must remain only within the secure enclave for some specified purpose. In effect, our conceptualization of redaction in data flows within our process is the same. In the context of reconciling conflicts across data flow policies, redaction can be performed on any datum which is itself a collection of subsumed data. In our ontology, this means the collection is a superclass. This is because the act of redacting data eliminates one or more subsumption relationships between the collection datum itself, and the subsumed data types. Redaction results in a new, redacted datum which is fit for a more general low-confidentiality purpose, as compared to the original datum which was only suitable for a specific high-confidentiality purpose.

As per our running example, Figure 4 shows a basic profile that has been retrofitted to permit Zumwalt to share radar data with a third party, which is a friendly fleet. The definition for radar data in this instance refers to data about friendly vessels, data about enemy vessels, and data about terrain/surface objects, which is expressed through the subsumption relationship seen on the left side of Figure 4. Internally, Zumwalt collects this radar data within its own secure enclave from the vessel’s onboard radar system, and this data is intended for consumption or transfer with an unspecified high confidentiality purpose only. It should be noted that in this scenario, the policies which govern consumption of this data internal to the Zumwalt are not within the scope of this discussion, as they are not related to the risk of data spill during transfer to a third party, and have been left out of the profile for simplicity. The area of security interest for our analysis in this scenario is with
respects to the requirements that govern how this data may be transferred to third parties outside of Zumwalt’s secure enclave, where the risk of data spill does exist.

**Figure 4. Simple Radar Data Sharing Profile for Zumwalt**

Next, we break down the individual rules that are seen in the profile introduced in Figure 3. The conflict is highlighted in Figure 4. Here we unravel the conflict and show how the redaction mechanism can be applied. First, we show the English language interpretation of a rule. Below, we show the application profile language used to express this rule, and the corresponding formalization in DL which is generated by our parser.

1. Permit collection of collected radar data from Zumwalt’s radar system, designating it as high-confidentiality data.

<table>
<thead>
<tr>
<th>Application Profile Language</th>
<th>Formalization in Description Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P COLLECT collected_radar_data FROM radar_system FOR high_confidentiality</td>
<td>$T = p0 \equiv \text{COLLECT} \land \exists \text{hasObject. collected_radar_data}\land \exists \text{hasSource. radar_system}\land \exists \text{hasPurpose. high_confidentiality}$</td>
</tr>
</tbody>
</table>

2. Permit transfer of data about enemy vessels to friendly fleet members for general, low-confidentiality purposes.

<table>
<thead>
<tr>
<th>Application Profile Language</th>
<th>Formalization in Description Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P TRANSFER enemy_data TO friendly_fleet FOR low_confidentiality</td>
<td>$T = p1 \equiv \text{TRANSFER} \land \exists \text{hasObject. enemy_data}\land \exists \text{hasTarget. radar_system}\land \exists \text{hasPurpose. low_confidentiality}$</td>
</tr>
</tbody>
</table>
3. Permit transfer of all collected radar data to friendly fleet members for general, low confidentiality purposes. This rule generates a conflict, which is explained below.

<table>
<thead>
<tr>
<th>Application Profile Language</th>
<th>Formalization in Description Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P TRANSFER collected_radar_data TO friendly_fleet FOR low_confidentiality</td>
<td>T \equiv ; p \in \text{TRANSFER} \land \exists \text{hasObject.} \text{collected_radar_data} \land \exists \text{hasTarget.} \text{friendly_fleet} \land \exists \text{hasPurpose.} \text{low_confidentiality}</td>
</tr>
</tbody>
</table>

4. Permit transfer of data about friendly vessels to friendly fleet members for specific, high-confidence purposes.

<table>
<thead>
<tr>
<th>Application Profile Language</th>
<th>Formalization in Description Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P TRANSFER friendly_data TO friendly_fleet FOR high_confidentiality</td>
<td>T \equiv ; p \in \text{TRANSFER} \land \exists \text{hasObject.} friendly_data \land \exists \text{hasTarget.} \text{friendly_fleet} \land \exists \text{hasPurpose.} \text{high_confidentiality}</td>
</tr>
</tbody>
</table>

5. Prohibit transfer of friendly fleet data to anyone for general, low confidentiality purposes. This rule conflicts with Rule 3, explained below.

<table>
<thead>
<tr>
<th>Application Profile Language</th>
<th>Formalization in Description Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>R TRANSFER friendly_data TO anyone FOR low_confidentiality</td>
<td>T \equiv ; p \in \text{TRANSFER} \land \exists \text{hasObject.} \text{collected_radar_data} \land \exists \text{hasTarget.} \text{Actor} \land \exists \text{hasPurpose.} \text{low_confidentiality}</td>
</tr>
</tbody>
</table>

Rule 5 works to prevent information about the friendly fleet leaking to third parties. This implies that regardless of the target for the data flow, a high-confidence purpose must be specified to justify friendly fleet data leaving the secure enclave of the Zumwalt. This requirement is instantiating a normal compartmentalization strategy for managing the flow of information, but this rule is in conflict with the intended new functionality. The retrofitted functionality requires the collected radar data to be shared with friendly fleet members, since collected radar data has been defined as friendly fleet data, as well as enemy fleet data and terrain data. This conflict can be resolved through redaction, as the prohibition only restricts the transfer of friendly fleet data for general, low-confidence purposes. By redacting this datum from its relationship with the collected radar data concept (given that it is the superclass for the other three types of data specified), we can define a new datum. Redacted radar data only contains enemy fleet data and terrain data, which are both unrestricted by purpose to high-confidence. The redacted datum may instead be used for low-confidence purposes as permitted by Rule 2), or implicitly permitted due to there being no rule which exists in this profile’s policy that governs the flow of terrain data.

Thus, the redaction statement syntax appears as follows:

<table>
<thead>
<tr>
<th>Application Profile Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>P REDACT(collected_radar_data -&gt; redacted_radar_data, friendly_data, low_confidentiality)</td>
</tr>
</tbody>
</table>

The above syntax represents the original datum to be redacted (linked with the -> operator to the new datum definition), the concept in the subsumption relationship to be removed from the original datum, and the newly established purpose for the redacted datum. This then permits the modified Rule 3 to read as follows, after resolving the conflict:
3. Permit transfer of all redacted radar data to friendly fleet members for general, low confidentiality purposes.

<table>
<thead>
<tr>
<th>Application Profile Language</th>
<th>Formalization in Description Logic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P TRANSFER redacted_radar_data TO friendly_fleet FOR low_confidentiality</td>
<td>[ T \models p_2 \equiv \text{TRANSFER} \land \exists \text{hasObject. redacted_radar_data} \land \exists \text{hasTarget. friendly_fleet} \land \exists \text{hasPurpose. low_confidentiality} ]</td>
</tr>
</tbody>
</table>

There is an alternative conflict resolution approach which is functionally the same as redaction, and may be more or less desirable depending on the functional context of the requirements. If there had been a high-confidentiality purpose specified for sharing the friendly fleet data in Rule 3, then the conflict would not have existed since Rule 5 only restricts the transfer of this datum for general low-confidentiality purposes. By narrowing the allowable purposes for which the data can be transferred and therefore consumed by a third party, the policy expressed in the profile would have been in alignment with the new requirements to transfer the radar data. However, there is a strong likelihood that the narrower range of purposes permissible for this datum may not be general enough for the third party’s intended purpose, especially if they had expressed a broader (low-confidentiality) intended purpose for the subsumed data.

A simple scenario which illustrates this case may be that the friendly fleet’s policy had intended, by their system’s internal design, to communicate the location of the enemy fleet to nearby civilians somehow. This action would still violate Rule 3 with respect to Zumwalt’s requirements since the enemy fleet data also includes collected radar data, and is therefore subject to being restricted to high-confidentiality purposes only, as low-confidentiality purposes are prohibited for transferring data. Due to the policy expression having deliberately chosen the more inclusive datum concept of all collected radar data, the previous resolution is to redact the datum and remove the general relationship with friendly fleet data.

The functionally similar alternative reconciliation is to completely recreate new policies which govern the data differently and express prohibitions and permissions uniquely for each individual datum concept. This secondary approach, while functionally the same as the action of redacting the original datum, may result in lengthy policies which do not gain expressiveness over the simpler act of redaction, and does not require completely redefining the data entities that are expressed in the original profile.
Generalization refers to the process of capturing a broader notion of a concept using inferences from a series of specific cases. In this context, generalization refers to the process by which a datum is transformed to represent a far broader notion of the original information, retaining enough precision to be useful, but not supplying enough precision to infer the original data. Examples in common usage include generalizing a position into a region, rather than a series of coordinates (such as Pacific Ocean, which is far more general, versus 0.0000° S, 160.0000° W, which is far less general) in order to obscure the precise location of a fleet by generalizing it to include a much broader area. The level of specified measurement precision selected for generalization must be appropriate to the datum. The act of performing generalization differs from redaction since the original ungeneralized datum cannot be a collection, or the generalization must be performed on all members in the collection. Generalization may also not have the same subsumed relationships as those applicable to the data undergoing redaction.

Generalization uses the same conflict resolution mechanism as redaction, in that it acts to realign the specified purposes for which a datum may be collected and transferred (or collected and used). For example, friendly fleet data such as locations may be generalized to express only regions rather than precise coordinates, in order to permit usage under low-confidentiality purposes among the collection of other data, as seen in Figure 4. In doing so, the superclass of what constitutes collected radar data is modified, and all subclasses in the collection which have been expressed already in the profile become permissible for use with low-confidentiality purposes. This is because friendly fleet data, enemy fleet data, and terrain data have all been generalized under this reconciliation strategy; each datum that is part of the superclass has been transformed and redefined as a new datum. These new data are implicitly permitted for transfer under general low-confidentiality purposes, just as terrain data had already been permitted without modification, as seen in Figure 4 and Figure 5.
Merging: Undoing Redaction and Generalization

Aside from simple differences in interpretation of policies and the conflicts that may result, our findings show that designers may accidentally introduce conflicts through actions that transfer data with mismatching purposes to those of their third party collector (or vice-versa) in order to realize system requirements that do not obviously conflict. This may occur in instances where the original design of a system did not consider collections of data to be usable as separate data when removed from the collection, as in our running example with collected radar data. In this example, the nature of the separate data in the collection was obscured by the subsumption relationship. The purpose of data which is combined from separate data can also be obscured. Merging disjoint data for more general purposes can create a new single datum that should only be used for the same generality of purpose as the original separate data. If this data is used for a more general purpose than any of those originally specified for any of the recombinant parts, there is a risk of repurposing data in a way that violates the original intention of the policy, which can lead to data spills. In our methodology, we define this act as merging, which requires two or more data to be used in conjunction with one another to create a new datum.

Given the transitive nature of data, and the inability to reason about data which moves past the designed system boundary of the protected enclave, proactive security assumptions must be made such that possible recombination of data may occur in any instance in which data has been transferred outside of the enclave for a general low-confidentiality purpose. This transfer and its associated risk of recombination with other data holds that the data may possibly be used for any similarly low-confidentiality purpose in conjunction with all other data of the same level of purpose. Conversely, if a datum is transferred for a specific high-confidentiality purpose, it may not be used for a low-confidentiality purpose without violating a policy or generating a conflicting requirement that must be reconciled.

The merge act combines one or more related data to create a new datum. This new datum may be equivalent to a datum originally specified for a high-confidentiality purpose, but had been previously redacted. The outcome of a merge act may yield more information than the original high-confidentiality purpose of a constituent data point. For example, a datum describing the position of a destroyer fleet at a time \( T \) only has the power to assert that the fleet was at that position at that time, which may be sufficient for some specific low-confidentiality purpose, such as establishing a point of rendezvous during a fleet maneuver. However, possessing additional data about the position of a fleet at time \( T + 1 \), \( T + 2 \) and so on can yield sufficient information to determine the previous bearing of the fleet and/or the likely future course of the fleet. This additional information may be restricted to a higher confidentiality purpose than was unintended for the data by its original specification, which could serve to mitigate the likelihood of merging this data without a specified high-confidentiality purpose. Generalization of this data may also serve as a countermeasure for this merging, since less precise coordinate data at all of these times would render an adversary unable to track the fleet precisely enough for the data to be useful. Increasing the amount of data points that are merged together in this basic example can further increase the precision or predictive power for a determination of the fleet’s movements, even in the presence of course corrections or evasive maneuvers which attempt to obscure the true intent of the fleet. This may also counteract the generalization strategy if some data points have been generalized while others have not, because the data flows transferring generalized data versus the ungeneralized data had low- and high-confidentiality purposes specified, respectively.
The same reasoning applies to other compositions of data with different information, even including the data that are not time sensitive. For example, determining the individual composition and outfitting of an individual ship in a fleet may yield only some tactical information about how Zumwalt’s executive officer may assay it as a single threat in isolation. However, finding similar data about all ships in a fleet and combining this information into a newly merged report about the whole fleet configuration may yield important clues about the strategic significance and intent of the fleet, far beyond the original tactical information of any one given ship’s threat.

In order to shield against an adversary performing this same merging of disjoint information to their advantage, the DoD must safeguard against the likelihood that data flows with specific purposes are repurposed for more general, low-confidentiality purposes. Otherwise, there is a risk that similarly purposed data flows may be merged.

Merge-related conflicts may be detected by determining the generality of the purpose in which data is transferred outside of the secure enclave and to a third party, which is a typical IA strategy. One general principle of IA strategy is that information should be disseminated on a need-to-know basis (based on the purpose of its usage, as per Bell-LaPadula), and those who do not need to know will not receive the data in question. Tracing data flows and matching the generality of the specified purpose is key to determine if a datum may be used for a different purpose elsewhere. Recall the example seen in Figure 5: If the friendly fleet had seen fit to recompose the original, non-redacted radar data using the redacted radar data in conjunction with the terrain data and enemy fleet data that they received, Zumwalt’s data sharing policy would be in violation due to the friendly fleet’s merge action.

Sending data for a specific, high-confidentiality purpose means that the collection actions that correspond to a datum’s transfer from the perspective of a third party must also match our high-confidentiality purpose, as must all subsumed classes of data related to the datum in question. If there is a mismatch in purpose, then there is a risk that any of the subsumed data may be repurposed for a more general purpose, which then increases the risk that it will be recombined with other subsumed data for the same superclass. This mismatch is in itself a conflict, as it violates the specified purposes expressed in the profile, and can be identified in the model by using our conflict detection methodology. This pinpoints the specific policies that are in conflict, and these policies become the subject of our conflict reconciliation strategies.

**Evaluation and Identifying Recertification Triggers**

Based on our running example, we surmise that conflicts may arise as a result of unintended or implied actions in a profile designer’s expression of data as a collection, which we see as a direct result of the new requirements imposed by new system functionality. Designers may express policies without having considered the individual elements in collections and the implications of adding new features that require usage at a different level of purpose than the original specification. In our running example, the original definition of collected radar data and the prohibition of sharing any data about friendly vessels implicitly did not consider that some future requirement would need this data to be shared with a third party. This conflict arises as a result of developing profiles that meet current system IA requirements, as well as functional requirements, but these profiles may not be useful as functional requirements change. Over time, operational requirements change, system designs change, and each of these changes may be unpredictable but still necessitates re-evaluation. Therefore, we must evaluate the effectiveness of our approach in terms of conflicts that can be identified and reconciled based on these changes.
We have identified that conflicts may result from overlapping organizational granularities expressed on one profile, and these conflicts must be reconciled. We have also identified that conflicts may result from changes in the intended functionality for a system governed by its profile’s policy. Recertification is then triggered by the reconciliation of these conflicts because they necessarily result in further changes to the profile beyond those originally introduced by new functionality. We assert that the recertification triggers are those actions that change a system’s profile sufficiently to merit reanalysis of the policies expressed. These actions include (but are not limited to) the addition of new features, modification of an existing behavior or feature requiring new connectors to outside parties, and modifying an existing connection to an outside party to serve a new purpose.

Our running example suggests that adding a new feature to a system is likely the most obvious recertification trigger, as it is functionally the same as modifying existing behaviors, and may also involve new connectors to outside systems for which data flows must be analyzed. In the case of Zumwalt introducing new functionality for its radar and information sharing systems, we could see that conflicts arose because of the mismatch in intended purposes for the previous policy which prohibited sharing of data about friendly vessels. This conflicted with the intended new functionality to share radar data with a friendly fleet, because the original policy was not intended for this functionality given the collection of friendly, enemy, and terrain data within collected radar data. Under circumstances such as these, the conflicts can be reconciled using the strategies we have defined, and profiles can be rapidly re-evaluated using our tool. Any new conflicts that arise may be reconciled using the same strategies until all of the conflicts are eliminated. Under this approach, potential conflicts that lead to data spill risks can be reasoned about prior to the development of the components which act as recertification triggers, or may be remediated if they are found after they have been instantiated as part of a new design or upgrade.

**Scalability Evaluation**

We must also evaluate our approach in terms of its efficient application and scalability to increasingly large profiles. In order to determine whether this approach is computationally scalable for extremely large compositions of systems, a performance simulation was conducted. This simulation was used to determine how much time was required for the language and DL reasoning engine to reason over profiles and detect conflicts, based on the size of the profiles. The simulation was designed to examine a much larger number of requirement statements than the simplified 5-rule profiles used in our running example, and strived to simulate profiles that would be as large as existing systems’ requirements. Previous case studies showed that the number of rules in a policy that we would expect to see in an integrated service scenario in commercial civilian applications is approximately 144 statements, so this provided a basis for determining the extent to which our simulation should scale, but no relative measure for performance.

We held the number of concepts and individuals constant and varied the number of rules to determine the time required to reason over the entire profile as the size of the profile increased. The scalability of our approach is largely dependent on the DL inference engine used to analyze the OWL ontology, rather than the tools which parse the Application Profile Language syntax. For the purposes of our study, we used the HermiT reasoning engine, as previous studies had shown it to have the fastest performance (Breaux et al., 2014) compared to other contemporaries (such as Pellet, Fact++, and Racer).

Our simulation results suggest that the parsing process, reasoning process, and output occurs expeditiously enough to claim that it scales quasi-linearly to analysis of profiles involving hundreds of data flows, and hundreds of rules. Our simulations show that even the
largest profiles within our test dataset could be analyzed in under 400 seconds. We expect that this would be sufficiently little time to permit modification and reanalysis of similar profiles several times in rapid succession. However, with no point of comparison with other tools that perform this functionality, we have no objective basis for which to determine the impact of our performance claims. We can only conjecture that for most purposes, it would be sufficient to be able to reanalyze a real profile of corresponding size to our largest simulated profiles three times within one hour, and use this as a baseline for our performance analysis. This conjecture is based on the fact that there is substantial cognitive load on the analyst to determine the correct conflict reconciliation approach for the given conflicts they encounter. However, these claims may be substantiated better in controlled experiments that quantify the amount of time required to instantiate each conflict resolution strategy, in order to validate this conjecture and substantiate the overall impact of our performance claim. An exploration of similar methodologies may also yield a point of comparison for the relative speed of our approach.

Simulations were performed in groups of 27 repetitions, in which an ontology was randomly generated as a result of a syntactically accurate minimal profile that expresses generic data definitions and purposes. This profile would have no semantic meaning to an analyst since it does not express concepts within the true problem space of IA requirements, but it is valuable for performance analysis and simulations of the reasoning process since it is structurally similar to real profiles and can be easily varied in size. For profiles of this kind, there appears to be a proportionally increasing probability of a conflict arising to the number of requirement statements and individuals specified in each profile. We found that there is a relationship of approximately 1.13 conflicts found for each increment in the number of rules expressed in the simulated profile for profiles containing greater than 15 rules. This relationship is visible in Figure 8, which is a scatterplot that shows that the number of conflicts increases quasi-linearly with respect to the size of the profile. There does not appear to be a direct correlation between the reasoning times required and the number of conflicts found within the reasoning process; using Pearson’s correlation, there is not a statistically significant relationship found with \( r(874) = .36, p > .05 \). Sixteen data types were used uniformly across all simulations, and simulation groups ran with increasing numbers of statements (an increase of two statements for odd numbered runs, and three statements for even numbered runs) from three statements up to 80 statements. The number of actors specified was random and increased proportionally to the number of statements, beginning from two with a maximum of 113, with mean of 59 and standard deviation of 24.

Figure 7 is a scatterplot graphing the increasing size of profiles versus the time required to reason over them. The figure shows that there is a proportional relationship between the time required to reason over the profile and the profile size (which is also the size of the OWL ontology), but aside from some outliers, the largest proportion of reasoning time required remained below 400 seconds even as the number of security requirements increased. Minor outliers appear to be a result of increased conflicts, and extreme outliers may be explained by aberrations in the time measurements resulting from changes in the proportion of processor time allocated to our test environment versus background processes on the same system, rather than the structure of the ontology or the speed of the tool itself. Parsing an entire profile requires less than one second on average. The major portion of the end-to-end processing time is devoted to HermiT’s automated reasoning and final output of the detected conflicts.
Note. Arrows represent the subsumption relationships between concepts seen in the ontology. This ontology corresponds to the policy seen in Figure 4’s profile. Mechanically reasoning over this ontology using a DL reasoning engine permits us to determine where permissions and prohibitions acting on the same (or subsumed) data conflict.

Figure 6. A Simple DL Ontology, Showing 5 Rules, Consisting of 4 Permissions (p), One Prohibition (r), and 3 Datum Definitions

In large profiles, there may be hundreds of conflicting requirements. In the presence of hundreds of rules, any collection/usage, or collection/transfer rule pair may increase the likelihood of a conflict. Each conflict must be identified and reconciled in order to mitigate the risk of a data spill. Without automation to find and analyze these conflicts automatically, there is a far more substantial risk that they will be overlooked by humans due to the effort and repetition required to analyze these policies manually. Manual analysis is a long and tedious process that quickly becomes intractable for humans even with small numbers of policies analyzed in isolation. This process necessitates mechanical analysis and automation, only possible through formal analysis using the supporting software tools. An example visualization of the complexity of profile ontologies can be seen in Figure 6, which visualizes five rules and three datum definitions, using the Protégé ontology visualization tool from Stanford University and the University of Manchester. This visualization is the same profile seen in our running example. In order to manually analyze this ontology without tool support, analysts would need to cross-reference each prohibition rule with each permission rule to determine if they acted on the same datum. Then, this subset would need to be examined to see if the prohibitions restricted an action that existed in one of the
permissions. That same analysis would need to be repeated for each subsumed concept that existed for each datum, and for each purpose.

**Table 1. Scalability Analysis Summary**

<table>
<thead>
<tr>
<th>Scalability Analysis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Profile Parsing Time</td>
<td>&lt;1 second</td>
</tr>
<tr>
<td>Largest Profile Size</td>
<td>80 rules</td>
</tr>
<tr>
<td>Longest Profile Processing Time</td>
<td>400 seconds</td>
</tr>
<tr>
<td>Average Conflicts per Statement</td>
<td>1.13</td>
</tr>
</tbody>
</table>

**Figure 7.** Plotting the Number of Security Requirements in the Profile vs. the Amount of Time (in Seconds) Required to Reason Over the Entire Profile (and Its Resultant DL Ontology)

**Figure 8.** Plotting the Number of Requirements Expressed vs. the Number of Conflicts Detected in a Simulated Profile

*Note.* There is a clear proportional relationship with few outliers.

**Conclusions**

In this paper, we recounted our methodology for performing rapid recertification tasks using formal analysis with our Application Profile Language. Our approach to automated
conflict detection was detailed. Based on this methodology, we identified three conflict reconciliation strategies that may be employed to resolve detected conflicts, and illustrated the process with a running example.

In our future work, we plan to extend the automation in our tool to provide automated recommendations to analysts for employing these conflict reconciliation strategies on existing profiles. We also plan to perform further performance analysis in order to objectively characterize the time savings gained by using this tool versus manual processes, and emerging formal analysis methods.

References


Panel 19. Factors Influencing the Selection of Contracting Strategies

<table>
<thead>
<tr>
<th>Thursday, May 14, 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>3:30 p.m. – 5:00 p.m.</td>
</tr>
<tr>
<td><strong>Chair: Elliott Branch</strong>, Deputy Assistant Secretary of the Navy (Acquisition and Procurement)</td>
</tr>
<tr>
<td><strong>Antecedents and Consequences of Federal Bid Protests</strong></td>
</tr>
<tr>
<td>Tim Hawkins, Western Kentucky University</td>
</tr>
<tr>
<td>Michael J. Gravier, Bryant University</td>
</tr>
<tr>
<td>E. Cory Yoder, NPS</td>
</tr>
<tr>
<td><strong>Avoiding Terminations, Single Offer Competition, and Costly Change Orders With Fixed-Price Contracts</strong></td>
</tr>
<tr>
<td>Andrew Hunter, CSIS</td>
</tr>
<tr>
<td>Gregory Sanders, CSIS</td>
</tr>
<tr>
<td>Alexander Lobkovsky Meitiv, CSIS</td>
</tr>
<tr>
<td><strong>Price Analysis on Commercial Item Purchases Within the Department of the Navy</strong></td>
</tr>
<tr>
<td>Janie Maddox, NPS</td>
</tr>
<tr>
<td>Paul Fox, NPS</td>
</tr>
<tr>
<td>Dr. Raluca Gera, NPS</td>
</tr>
</tbody>
</table>

**Elliott Branch**—is the deputy assistant secretary of the Navy (Acquisition and Procurement) in the Office of the Assistant Secretary of the Navy (Research, Development and Acquisition). He is the senior career civilian responsible for acquisition and contracting policy that governs the operation of the Navy’s worldwide, multibillion-dollar acquisition system. Branch is the principal civilian advisor to the Navy Acquisition Executive and serves as the Department of the Navy’s Competition Advocate General for procurement matters and is the community leader of the Navy’s contracting workforce. Prior to joining the Navy Acquisition Executive staff, Branch was the first civilian director of contracts at the Naval Sea Systems Command. In that role, he led one of the largest and most complex procurement organizations in the federal government. As the senior civilian for contracting at NAVSEA, Branch was responsible for the contractual oversight of the nation’s most complex shipbuilding and weapons systems procurement programs. His duties involved the obligation and expenditure of approximately $25 billion annually. He is a member of the Senior Executive Service (SES). Members of the SES serve in the key positions just below the top presidential appointees. They are the major link between these appointees and the rest of the federal workforce. SES members operate and oversee nearly every government activity in approximately 75 agencies. Branch spent time in the private sector, where he specialized in acquisition and project management education, training and consulting for the federal workforce and its associated contractors. In this role, Branch was responsible for the design, development, delivery and maintenance for a wide variety of course material ranging from project management to contract law. Branch’s clients included Computer Sciences Corporation, QSS Group, BAE Systems, the Pension Benefit Guaranty Corporation, and the Departments of Defense, Energy, Justice and State. Prior to that, he served as the chief procurement officer for the Government of the District of Columbia, where he was the agency head responsible for procurement operations, policy, and for formulating legislative proposals for local and congressional consideration. Branch led a staff of over 200 employees that supported over 40 city agencies, administered a $14 million annual operating budget, and oversaw the
placement of $1.5 billion, annually, in city contracts. Before joining the District government, Branch held various positions in the SES with the Department of the Navy (DoN). In 1993, he became a member of the SES as the director, Shipbuilding Contracts Division, at NAVSEA. He next served as executive director, Acquisition and Business Management for DoN, responsible for policy and oversight of contract operations throughout the entire Navy. While in this position, he also served as project executive officer, Acquisition Related Business Systems. In this role, he was responsible for the formulation and execution of a multi-year effort transforming the Navy’s acquisition system from a paper-based system into one that made use of electronic technologies and methods. In this role, Branch was directly responsible for a portfolio of projects worth more than $200 million. Branch graduated with a Bachelor of Science degree in economics from the University of Pennsylvania and completed the Executive Program at the University of Virginia Darden School. He has received the Navy Distinguished Civilian Service Medal, the David Packard Excellence in Acquisition Award, two Presidential Rank Awards for Meritorious Executive, and the Vice Presidential Hammer Award for Reinventing Government.
Antecedents and Consequences of Federal Bid Protests


Michael J. Gravier, PhD, CTL—is an associate professor of marketing and global supply chain management at Bryant University. He received his PhD in marketing and logistics from the University of North Texas, MS in logistics management (with a specialization in transportation management) from the Air Force Institute of Technology, and BA with majors in Spanish and anthropology from Washington University in St. Louis. Prior to his academic career, he spent 12 years as an active duty logistics readiness officer in the U.S. Air Force and worked as a research assistant in the Washington University School of Medicine. He has research interests in supply chain ethics, procurement, logistics pedagogy, logistics in emerging countries, and the evolution of supply chain networks in response to risk factors like obsolescence and changing information needs. [mgravier@bryant.edu]

E. Cory Yoder, CDR, USN (Ret.)—is a senior lecturer and academic associate at the Naval Postgraduate School. He holds an MA in National Security and Strategic Studies (Naval War College); an MS in management (Naval Postgraduate School); a BS in business management (Indiana University); and a Business Resource Management certificate (University of Virginia–Darden), and is a Beta Gamma Sigma honor society member and RADM John J. Schieffelin awardee. He has published in, but not limited to, the Journal of Public Procurement, Army AL&T, NATO Proceedings, and Defense Acquisition Review Journal. His current research interests include strategic contract planning for operations, humanitarian and disaster contract response, performance and compliance auditing for contracting organizations, and creating business efficiencies through sound contracting and business disciplines. [ecyoder@nps.edu]

Abstract

The fear of receiving a bid protest is said to affect acquisition strategies, yet it has not been empirically explored. Based on the Public Value Framework and interviews with contracting personnel, this research tests a model of antecedents to and consequences of the fear of a protest. Survey data was obtained from a sample of 350 contracting personnel. The fear of protest is mitigated by having sufficient procurement lead time and by source selection competence, and increased by protest risk. Fear of protest increases compromised technical evaluations, added procurement lead time, and transaction costs, while it decreases contracting officer authority and is associated with source selection method inappropriateness. Compromised technical evaluations, in turn, decrease contractor performance while contracting officer authority increases contractor performance. Thus, findings suggest that fear of protests affect acquisition strategy decisions, which, in turn, affect contractor performance. The research concludes with several managerial implications, study limitations, and future research directions.

Research Issue

Seemingly, the rate of errors in acquisition procedures is increasing since the quantity of bid protests received each year is increasing. In 2009, 1,989 protest cases were filed across the federal government. Of those protest cases that made it to a decision (i.e.,
the few that were not dismissed, settled, or withdrawn), only 18% were sustained, but 45% of all protest cases were effective (either sustained or resulted in corrective action by the agency prior to decision). In 2013, 2,429 protest cases were filed; 17% were sustained, and 43% were effective. When accounting for the number of contract actions awarded (i.e., the protest opportunity), the number of protests increased from 2011 to 2013 (from 0.014% to 0.018% of contract actions including delivery orders).

Bid protests come with monetary and non-monetary costs. An agency must incur costs to prevent a potential bid protest (e.g., thoroughly documenting and substantiating proposal evaluations and trade-off decisions), to defend against an actual protest lodged, and to take corrective actions. The end users bear costs as well, since their requirements are delayed or go unfulfilled.

The Department of Defense (DoD) acquisition workforce seemingly believes that it is important to avoid protests. This desire to avoid a protest is the driving force behind acquisition decisions, internal and external policies, and resources applied to mitigate the threat of a protest. Evidence suggests that agencies sometimes change their acquisition strategies due to fear of protests. For example, fear of a protest could prompt officials to try to structure a contract in a manner they deem less likely to be protested, such as using a lowest-price, technically-acceptable (LPTA) source selection method instead of a full tradeoff (Schwartz et al., 2013). Other reactions include awarding more contracts than intended to avoid a protest (e.g., Littoral Combat Ship). While scholars and the Government Accountability Office (GAO) have identified these deleterious effects of bid protests on the government (Gordon, 2013), no research to date has quantified them. Specifically, we do not know the magnitude of fear of protests. Neither do we know the extent that fear of a protest affects acquisition strategies nor the lengths that acquisition professionals will go to avoid a protest.

Purpose and Research Questions

The purpose of this research is to quantify the magnitude of protest fear, and to explore the antecedents and consequences of protest fear. The research questions addressed include the following:

- Do bid protests lead to sub-optimal acquisition strategy decisions?
- Do bid protests affect source selections?
- If yes, does contractor performance suffer?
- How are contracting officers’ authorities affected?

Theoretical Frameworks

Fear of a protest is understandable. Significant time is consumed addressing a protest. High dollar contracts, in particular, hold great interest to media and elected officials. A protest would reflect negatively on the contracting official as well as the contracting and program offices. There may be an element of shame if a source selection is protested, particularly if there is a notion that management would not support the contracting officers and that the protest may reflect poorly on them. With these concerns in the back of a contracting officer’s mind, there can be a tendency to take measures in order to avoid a protest that can sub-optimize source selection decisions and outcomes. For example, the contracting officer may rely too heavily on the LPTA method rather than utilizing a full tradeoff approach. The motivations and reactions to fear of protest can be better understood by applying the public value framework (PVF).
**Public Value Framework**

PVF was introduced by Harvard professor Mark H. Moore and has been used to evaluate and identify value in, mainly, the public sector. Value in the public sector is much different than it is in the private sector. Often in the private sector, industry uses shareholder value as a means of evaluating itself. The public sector, however, is much different. The PVF has been utilized to “get public managers thinking about what is most valuable in the service that they run and to consider how effective management can make the service the best that it can be” (Coats & Passmore, 2008, p. 4).

The PVF can be explained by the strategic triangle (Heymann, 1987; Moore, 1995). The three elements are public value, legitimacy support, and operational capability. In contrast to private sector operations, the government’s strategy does not revolve around a specific bottom line, such as shareholder wealth. Contracting professionals are often satisfying multiple stakeholders such as regulatory requirements (e.g., the FAR), internal customers, the private sector, and the taxpaying public. The first element, public value, “directs managerial attention to the value proposition that guides the organization. For an enterprise to succeed in producing value, the leaders of the enterprise have to have a story, or an account, of what value or purposes that the organization is pursuing. They need a reason for the organization’s existence, a claim about the way in which the world would be made better through the operations of the enterprise” (Moore, 2000, p. 197). In essence, value in a governmental organization equates to mission. Contracting officers add value by connecting capable suppliers to internal organizations in need of quality goods and services.

Legitimacy and support “directs managerial attention to the question of where the support for pursuing the value will come from. It is not enough that an entrepreneurial leader judges some purposes to be valuable. Others, who provide the necessary financial resources and authorization, have to agree with that judgment. In government, those others include citizens, elected representatives, interest groups, and the media, which has been called the “authorizing environment” of the organization (Moore, 2000, p. 198).

Finally, operational capacity focuses attention on the question of whether sufficient know-how and capability exist to achieve the desired results. Often, this capability lies entirely in the organization that the manager leads. However, sometimes it lies outside the organization’s boundary, and the organization has to find ways to engage capacities beyond its own to achieve the desired result by creating partnerships of various kinds. (Moore, 2000, p. 198)

Contracting professionals add value by helping to meet the operational needs of the government and, at the same time, provide fairness and address the various public policy issues that are required by law and regulation. When these align, customers receive what they require at a fair and reasonable price, and this satisfies the requirements of governing policies. Through this, government contracting professionals add value to all of their stakeholders. Contracting officers sometimes take steps throughout the acquisition process to avoid a protest, such as minimizing discussions or even employing an LPTA source selection process when a full tradeoff method is more appropriate (Gordon, 2013, pp. 36–37). When this occurs, the contracting system is not optimizing its value.

Contracting officers are also accountable to provide fairness to commercial entities with which they contract for goods and services. Often, though, a fear of a bid protest will result in awarding more contracts than would have been awarded if there was no fear of a bid protest. In multiple-award contracts, there is a minimum dollar value that the government is obligated to pay (as consideration). This results in increased spending of taxpayer money...
that could have been more efficiently spent by awarding to fewer, more competitive contractors. For example, the DoD Inspector General (DoDIG; 2009a) found that under the Navy’s Seaport-E program, the Navy awarded 1,279 contracts for professional services, yet 975 (75.6%) never received a task order. Each of these contracts required either a $10,000 or $2,500 minimum obligation. Added contracts also create extra work for the contracting officer to administer, duplicate inventory, can increase transportation costs, result in non-optimal use of taxpayer money, and often upset contractors who never get an award under a multiple award contract. Although a reduced risk of a protest is accomplished, ultimately less value is added by the contracting process. What this does not accomplish is a best option for the customer or the taxpayer, nor does it provide fairness to the stronger contractors.

**Antecedents to Fear of Protest**

**Sufficiency of Planned Procurement Administrative Lead Time:**

Sufficient planned procurement administrative lead time (PALT) represents the extent to which adequate time is allotted to accomplish a source selection. Insufficient PALT is often the result of funding constraints that occur toward the end of the fiscal year. Expedited requirements and poor planning are common reasons that can lead to insufficient PALT. Failure to allocate sufficient lead time to properly define requirements (Hawkins et al., 2011)—evaluation criteria, and instructions to offerors; train the technical evaluators; evaluate proposals; document evaluations and tradeoffs; and prepare for and brief decision-makers—makes protestable errors more likely to occur. Sufficient time bolsters acquisition team capability to perform a source selection: “Time has become a major variable in the typical buyer’s decision process of choosing a supplier” (Hansen, 2009, p. 234).

PVF’s operational capability experiences a positive relationship with sufficient procurement time. Time affords the ability for acquisition teams to apply their knowledge and skills; absent sufficient time, operational capability is constrained. Therefore, it is posited that

\[ H1: \text{Insufficient PALT is positively related to fear of protest.} \]

**Contracting Officer Competence:**

PVF holds that operational capability is necessary in order for government activities to deliver value. Operational capability represents the requisite knowledge, skills, and abilities—all of which develops with experience. The more experience a contracting officer has, the less concern of a protest there should be since the individual has acquired more knowledge in techniques and practices to prevent bid protests and prevail in the event of a protest. Buyer experience has been found to affect government procurement processes (Hawkins & Muir, 2014).

Time spent in a competency correlates strongly with self-reported proficiency levels in that competency (Federal Acquisition Institute [FAI], 2012). Econom (2006) argued that federal agencies must consider contract management as a core competency because the functions performed by third-party contractors are often essential in successfully achieving organizational goals. She concluded that the success of acquisition organizations is largely dependent on hiring personnel who possess the right mix of skills, abilities, experience, and training. Other studies have also found that the right mix of experience and competency is critical to achieving contract performance outcomes (United States Merit Systems Protection Board, 2005). Therefore, it is hypothesized that

\[ H2: \text{The greater a contracting officer’s competence, the lower the fear of protest.} \]
Consequences of Fear of Protest

Compromised Technical Evaluation: Evaluation factors and significant sub-factors must (1) represent the key areas of importance and emphasis to be considered in the source selection decision; and (2) support meaningful comparison and discrimination between and among competing proposals (Federal Acquisition Regulation [FAR], 15.304(b)). Agencies must evaluate the proposals and assess their relative qualities based only on the factors and sub-factors specified in the request for proposal (RFP; Rumbaugh, 2010). Deviations from the strict language defining the meaning of factors and sub-factors can invite protests. Technical evaluators often do not understand or appreciate this constraint. In the PVF terms, poorly trained or technical evaluators unknowledgeable in source selections inhibit the agency’s operational capability. Exacerbating this problem are cases in which evaluators assessing proposals are not the same individuals who defined the meaning of the factors and sub-factors, leading technical evaluators to develop their own interpretation or agenda.

For these reasons, the contracting officer, legal advisor, and contracts committee advisors often require numerous, meticulous changes to precise wording of evaluations. Definitions of the factors or sub-factors may not account for meaningful distinctions, or evaluators are constrained on what they can say in the evaluation even though the point otherwise intended may make a meaningful distinction between offers. Additionally, often this phenomenon reflects a lack of foresight—sometimes preventable, sometimes not. Sometimes, only upon evaluation of proposals is the distinction illuminated. At this point, the source selection team must weigh a delay in the schedule with the benefit of changing the definition of factors or sub-factors to account for the meaningful distinction, and allow offerors time to revise their proposals. Often, however, the customer is not willing to delay the source selection, and the sub-factors are not revised. Therefore, it is posited that

H3: Fear of protest is positively related to compromised technical evaluations.

Source Selection Method Appropriateness:

Competitive formal source selections may follow one of several methods—lowest-price, technically-acceptable (LPTA), price-performance tradeoff (PPT), or a full tradeoff. According to FAR 15.101-2, the LPTA source selection process is appropriate when best value is expected to result from selection of the technically acceptable proposal with the lowest evaluated price. There are many reasons why a contracting officer might opt for the LPTA method. One major benefit of this strategy is that the agency can greatly shorten the evaluation process because once the low price proposal has been found to be technically acceptable, there is no need to evaluate the acceptability of any other proposals (Cibinic et al., 2011, p. 680). The source selection method appropriateness depends on the requirement and the buying situation. Generally, the greater the performance risk, criticality of the requirement’s successful delivery to the agency’s mission, dollar value, environmental dynamism, uncertain requirements, and complexity, the more important contractor performance becomes and the less critical cost/price become. In these cases, an agency may decide that the best-value offer is determined by a full tradeoff of price and non-price factors. A full tradeoff process is appropriate when it may be in the best interest of the government to consider an award to a company other than the lowest-priced offeror or other than the highest technically rated offeror (FAR 15.101-1).

But, agencies may not select the source selection method that is best suited to the requirement and the buying situation. Today’s budget-constrained environment may influence managers to prefer LPTA over a full tradeoff. Managers may also wish to avoid a protest, in which case the LPTA method is clearly the lower-risk alternative. In fact, Air Force
acquisition leaders, following several bid protests and failed attempts to acquire a new tanker aircraft, seriously considered an LPTA method for a multi-billion-dollar weapon system (Pocock, 2009). Finally, quite often managers prioritize the contract award date (i.e., PALT) over due diligence in contractor selection (Hawkins, 2012). Therefore, we posit that

\[ H4: \text{Fear of protest is negatively related to source selection method appropriateness.} \]

**Added PALT:**

Naturally, as the concern over a protest grows, acquisition teams take added measures to prevent them. This is often manifested in increased reviews resulting in increased iterations of source selection documents such as source selection plans, requests for proposals, technical evaluations, small business strategy, comparative analyses, briefing charts, source selection decision documents, and evaluation notices to offerors—just to name a few. These revisions consume time during the source selection. Additionally, a conservative stance may result in added rounds of discussions to clear up all proposal deficiencies and weaknesses—a concept referred to as *technical leveling*. Conservatism may also result in retaining otherwise non-competitive offerors in the competitive range, adding time to negotiate with and to evaluate another offer. Therefore, it is hypothesized that

\[ H5: \text{There is a direct positive relationship between fear of protest and the added PALT.} \]

**Contracting Officer Authority:**

Contracting officers uniquely hold authority to enter into, administer, and terminate contracts. They are the only individuals authorized to bind the U.S. government. Contracting officers are responsible for (1) ensuring that all the necessary actions for effective contracting are accomplished, (2) ensuring compliance with the terms of the contract, and (3) safeguarding the interests of the U.S. government in its contractual relationships. In terms of the Public Value Framework, the role of the contracting officer is to exercise his or her authority, thereby protecting the various stakeholders’ interests (e.g., taxpayer, contractor, government, internal customers). In this capacity, the contracting officer reinforces legitimacy support.

While contracting officers must request and consider the advice of specialists (e.g., law, engineering, finance, etc.), ultimately, decisions within their purview are their responsibility (FAR 1.602-2). Upon receipt of a protest, legal counsel must divert time and effort to defend the agency’s actions. Thus, legal counsel reviews the many iterations of the multitude of source selection documents to ensure legal sufficiency, compliance to regulations and policies, and to mitigate the risks of protests. With the consequences at stake, such as setting precedent, reputation, and invested time, legal counsel is typically conservative in attempting to prevent a bid protest. Since legal counsel brings their own unique legal authority and professional expertise, contracting officers and acquisition managers rely heavily on its opinions and recommendations. One interviewee shared, “We almost never move forward unless they [legal] give us their okay. It would be very, very hard—very challenging.” This comment alludes to the influence of legal counsel on acquisition and unit leaders. Contracting officers are likely to yield their decision-making discretion (e.g., removing an offeror from the competitive range) when legal counsel disagrees with them. Thus, legal counsel, in its advisory role, subtly, yet strongly, affects the contracting officer’s authority through its opinions and recommendations.

Other parties impose a similar phenomenon on contracting officers’ decision authority. For example, higher ranking contracts committee members and leaders may also
hold opinions on a particular source selection matter that are contrary to that of the contracting officer. In such cases, contracting officers may perceive unwritten career implications to making contrarian decisions. Thus, although certain statutory authority resides with the contracting officer, the reality is that such authority is yielded in practice. As protest risk—and thus, fear of a protest—grows, so does the involvement of legal counsel, other reviewing parties, and acquisition leaders. Increased involvement likely reduces the contracting officer’s perception of decision latitude. In some instances, contracting officers indicate that legal counsel would not allow them to make decisions—creating the organizational norm that legal has the final decision, not the contracting officer. Thus, we posit that

\[ H_6: \text{There is a negative relationship between fear of protest and the contracting officer's perceived authority.} \]

Transaction Costs:
The DoD has experienced a significant increase in the number of competitive source selection decisions which are protested by industry. Protests are extremely detrimental to the warfighter and the taxpayer. These protest actions consume vast amounts of the time of acquisition, legal, and requirements team members; delay program initiation and the delivery of capability. (Young, 2007, p. 1)

Transaction costs reflect the monetary costs of resources devoted to executing a formal source selection—largely comprised of labor costs of the different acquisition professionals involved (contracting officer, contracting specialist, technical evaluator, legal, cost/price analyst, past performance team, program manager, Small Business Administration representative, and consultants). Transaction costs could be considered an opportunity cost of resources not devoted to other work requirements (e.g., contract and program administration). As the risk of a protest increases, and the fear of a protest, more personnel are involved and they allocate more of their time and effort to defending against a potential bid protest. Thus,

\[ H_7: \text{There is a direct positive relationship between fear of protest and transaction costs.} \]

Contractor Performance:
A central aspect of the Public Value Framework is providing value through the organization’s mission. An organization’s mission is increasingly performed or supported via outsourced contracts. Thus, in order for the government to attain mission success, contractors must be successful. They must perform well under the obligations of their contract. The source selection process can affect the level of performance ultimately received.

When the government utilizes a best-value source selection method, technical evaluators apply evaluation factors and sub-factors to proposals to determine the best-value offer. This process helps the government to hedge against substandard and/or non-performance by weeding out the less-capable firms (or teams of firms). The premise of source selection is that by applying the evaluation factors and sub-factors, a very capable contractor has higher odds of being deemed the best-value offer. Nonetheless, the government struggles in its efforts to select and sufficiently define high-quality technical factors and sub-factors such that they can make meaningful distinctions between offers (Rumbaugh, 2010). Once weaknesses in evaluation factors are realized, particularly after
receipt of proposals, acquisition teams are reluctant to fix the factors by amending the RFP and inviting revised proposals since these actions delay the acquisition milestones. Additionally, conservative evaluators (and their advisors), for fear of protest, often engage in multiple rounds of discussions that essentially level the playing field of competitors, and often they retain mediocre offerors in the competitive range for fear of receiving a bid protest. Had the evaluation criteria been better able to distinguish amongst the firms, the propensity to retain mediocre firms within the competitive range would be diminished. Together, therefore, it is expected that

H8: There is a negative relationship between compromised technical evaluation and contractor performance.

Contracting officers are generally more cognizant of and empathetic to the effects of contractor performance on the requiring organization’s mission attainment than are risk-averse advisors—such as legal counsel. Thus, contracting officers may prioritize mission performance over protest risk in making key decisions during a source selection. For example, they may be more apt to remove a less-competitive or less-capable offeror from the competitive range, assign a lower proposal rating, and not engage in added rounds of discussions solely to mitigate protest risk (thereby mitigating technical leveling). In some cases, contracting officers may also be more apt to choose a full tradeoff rather than a LPTA source selection method as the most appropriate means to attain the best-value contractor. The full tradeoff method allows the government the flexibility to pay more for superior capability and/or past performance when warranted. But, this method also requires more effort, invites more error, and thus, protest risk, since the procedures are so nuanced. These actions reduce the odds of having to award a contract to a less-capable contractor, for example, in the case of a LPTA source selection method. Therefore, it is hypothesized that

H9: There is a positive relationship between source selection method appropriateness and contractor performance.

When an individual holds decision-making authority coupled with accountability for the results of decisions (e.g., a contracting officer), he or she tends not to defer decisions entirely to others. This is not to say, however, that others are not consulted. In public contracting, similar to input from advisors on source selection method choices, advisors review all of the written technical evaluations with respect to the evaluation criteria published in the request for proposals. They screen for errors, omissions, consistency, and other matters of compliance with laws, regulations, and policies in an effort to mitigate the odds of receiving a bid protest. In doing so, advisors often limit what the technical evaluators can say. Such scrutiny can make difficult the ability to meaningfully discriminate between proposals. Similar to the previously discussed rationale, while contracting officers also review the technical evaluations for errors, they are more apt to accept more risk. Thus, it is posited that

H10: Contracting officer authority is negatively related to compromised technical evaluations.

Taken together, then, it is expected that a contracting officer with decision-making authority—and who does not defer that authority to others—will make decisions that accept more risk yet does not impede the selection of the best-suited contractor for the task at hand. When the selection is not constrained by procedures, greater decision latitude results in a better match between the offeror’s capabilities and the contractual requirements. This better match should facilitate better performance. Examining the troubled U.S. defense acquisition system, the Defense Business Board concluded, “Of the eight findings, three of them concern the acquisition workforce, a large group of dedicated public servants who
work diligently, but ultimately struggle within a broken system that is focused on avoiding mistakes rather than producing more, in less time, at less cost” (Punaro, 2012). Therefore, it is posited that

\[ H11: \text{There is a positive relationship between contracting officer authority and contractor performance.} \]

The relationships posited above are visually depicted in Figure 1. Fear of protest can also be affected by the criticality of the requirement and by protest risk. Therefore, these constructs are shown as control variables.

**Figure 1. Conceptual Model of Fear of Protest**

**Methodology**

This study employed a mixed design (Creswell, 2003) of qualitative and quantitative analysis. The qualitative work involved discussions with academicians and interviews with practitioners to ensure face validity and construct validity, to construct and validate a conceptual model, and to develop survey items to measure the constructs—many of which did not previously exist. Next, the research employed structural equation models using cross-sectional survey data. The remainder of this section details the qualitative design, interviews, survey development, sample, data collection, and reliability and validity.

**Interviews**

Contracting officers at two military organizations were chosen for interviews due to (a) proximity, (b) a willingness to support the research, and (c) the availability of a wide variety of contract types and contracted goods and services for wide generalizability (e.g., external validity). A series of questions was asked to each participant (Appendix A).

Eighteen individuals were interviewed over two days. Demographics of each respondent can be found in Table 1. Each interview was recorded and transcribed. The average interview lasted 26 minutes. The interviews resulted transcribed into 229 pages. Informants were given a copy of the conceptual model during the interview and asked whether they agreed with the independent variables being used. They were also asked if they would add or omit any. One respondent stated, “Okay. This is good. I don’t see
anything that I need to add.” Another contracting officer stated, “I think this is a great research that you are doing because this is a bigger and bigger issue. I think you are right on.” Other statements that validated the model were, “I think I like the model. For the most part it says everything.”

Table 1. Informant Demographics

<table>
<thead>
<tr>
<th>Rank/Grade</th>
<th>Gender</th>
<th>Duty Title</th>
<th>Yrs. Exp</th>
<th># Source Selections (within 2 yrs)</th>
<th># Bid Protests (within 2 yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS-14</td>
<td>F</td>
<td>Supv. Contracting Officer</td>
<td>29</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>GS-13</td>
<td>M</td>
<td>Contracting Officer</td>
<td>26</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>GS-13</td>
<td>M</td>
<td>Contracting Officer</td>
<td>18</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>GS-14</td>
<td>M</td>
<td>Contracting Officer</td>
<td>36</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GS-13</td>
<td>F</td>
<td>Supv. Contracting Officer</td>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>GS-14</td>
<td>F</td>
<td>Contracting Officer</td>
<td>32</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>GS-12</td>
<td>M</td>
<td>Contracting Specialist</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GS-13</td>
<td>M</td>
<td>Contracting Officer</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>GS-12</td>
<td>M</td>
<td>Contracting Specialist</td>
<td>7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>GS-13</td>
<td>M</td>
<td>Contracting Officer</td>
<td>21</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GS-13</td>
<td>F</td>
<td>Contracting Officer</td>
<td>30</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>GS-13</td>
<td>M</td>
<td>Contracting Officer</td>
<td>22</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>GS-13</td>
<td>F</td>
<td>Contracting Officer</td>
<td>15</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>GS-13</td>
<td>M</td>
<td>Contracting Officer</td>
<td>26</td>
<td>20</td>
<td>1</td>
</tr>
</tbody>
</table>

Questionnaire Design and Construct Measurement

All scales measuring latent constructs used a Likert-type scale. Fear of protest is a term used for this research to identify the level of apprehension a contracting professional has about receiving a bid protest. No previously validated scales were available to measure the fear of protest; thus, scale items were developed from the interview data (Appendix B). The contracting officer authority construct described how empowered the contracting officer is to make final decisions during the source selection process. Similarly, there were no previously validated scales available for this construct. A three-question scale was used to measure the sufficiency of planned PALT in the milestones and allocated by the acquisition team and its managers to conduct the source selection (Hawkins & Muir, 2014). Compromised technical evaluation assessed the extent to which technical evaluators
complained about the limitations imposed on the wording of their written technical evaluations. Contractor performance is a measure of the contractor’s performance levels and the degree to which requirements were satisfied. The scale was adapted from Fawcett, Smith, and Cooper (1997), Cannon, Achrol, and Gundlach (2000), and Prahinski and Benton (2004). Source selection method appropriateness is the perceived extent that the chosen source selection method fits the requirement, the goals of the source selection, the commercial market, and the acquisition situation. There were no previously validated scales available for this construct. Source selection experience was measured as the number of source selections the respondent previously experienced. This could include FAR Part 15 (i.e., formal) and non–FAR Part 15 (e.g., simplified) source selections. It could also include those source selections to which the respondent served as the procuring contracting officer as well as those to which the respondent served as an advisor or reviewer. Added PALT objectively measured the difference between the planned PALT and the actual PALT. Transaction costs attempted to quantify the personnel costs based on amount of time spent on the source selection by each member of the team.

Survey Pretest

Six industry practitioners and academicians tested the initial survey. Feedback received was used to refine questions and limit survey length. As a result, one construct was removed, and the order of the survey questions was structured to reduce bias among scale items by mixing questions across constructs with like scales and scale anchors.

Sample

The population for this study consisted of civilian and military contracting personnel who had executed a FAR Part 15 formal source selection (i.e., a dollar amount greater than $150,000). This excluded simplified procurements that are generally less susceptible to bid protests. A list of e-mail addresses was generated using data extracted from the Federal Procurement Data System—Next Generation (FPDS-NG) database to encompass all transactions that fit the criteria previously stated.

The unit of analysis for this survey was a source selection. Since nearly all bid protests stem from a protestable action (e.g., a proposal rating, rating justification, or basis of a tradeoff analysis) associated with a source selection, this is the proper unit of analysis for the study. Respondents were instructed to answer the survey questions using their experience from their most recently completed FAR Part 15 source selection. The most recent source selection was required to serve as the basis of reference in order to prevent respondents’ self-selection bias.

Results

The fear of protest was empirically validated. To examine fear of protest, a structural equation model of its antecedents and consequences was tested and found to exhibit good fit to a sample of data from 350 FAR Part 15 source selections.

The less sufficient the planned procurement lead-time is thought to be, the level of fear of a protest increases. When acquisition personnel have less time than they believe is necessary to properly conduct the source selection, there are greater odds of making a mistake that could be protested. Additionally, a contracting officer’s competence—in terms of the number of source selections experienced—lowers the level of fear of bid protests.

As a result of protest fear, technical evaluations appear to be somewhat compromised. This is important since compromised technical evaluations also decrease contractor performance. While a fear of protest did not affect perceived source selection method appropriateness, protest fear was associated with the inappropriate use of the LPTA
source selection method. In turn, LPTA inappropriateness negatively affects contractor performance.

The fear of protest diminishes a contracting officer’s perceived authority (i.e., discretion in making decisions). This is important since diminished contracting officer perceived authority was found to decrease contractor performance directly. Contracting officer’s perceived authority also affects contractor performance indirectly by decreasing compromise technical evaluations.

The fear of protest is positively related to an increase in transaction costs. Costs were assessed in terms of the number of personnel involved in a source selection and their allocated time. The average cost per source selection was $235,236 (median = $165,832) with a standard deviation of $291,620. Notably, these costs are understated by considering direct salaries only; they exclude the fully burdened cost of a government employee. An average of 9 different people worked on a given source selection team in the various roles (an average of 3.5 full-time equivalents). Post hoc analysis showed that as the fear of protest increases, the number of personnel and the actual procurement lead time increase.

Implications

The more insufficient the planned procurement lead-time is thought to be, the level of fear of a protest increases. When acquisition personnel have less time than they believe is necessary to properly conduct the source selection, there are greater odds of making a mistake that could be protested. Interestingly, when they are rushed, contractual documents (e.g., statements of work) and pre-award communications (e.g., negotiations) could be compromised, which may, in turn, decrease contractor performance (Hawkins et al., 2011). Shortcuts could preclude the selection of the best contractor or result in selecting a contractor that does not fully understand the requirements. Thus, acquisitions should not be hastened short of the time thought to be adequate by the contracting officer. To prevent rushed acquisitions, standard lead times by type of source selection and by complexity of the requirement could be established.

A contracting officer’s competence—in terms of the number of source selections experienced—lowers the level of fear of bid protests. Therefore, efforts should be made to increase the number of source selections experienced by contracting officers. Of the 350 survey respondents, the average number of source selections experienced over a career was 36.7. That is just under 2.8 source selections per year. Note that this seemingly high number of source selections likely includes simplified buys and experience in a variety of roles such as a peer or committee reviewer as well as a contracting officer. There is no equal alternative to on-the-job-training (OJT), but source selection simulations and scenario-based training could be utilized as an alternative and as a supplement to OJT. If the acquisition community is relying solely on OJT, it can take a contracting officer and technical evaluators far too long to gain an adequate level of competence with FAR Part 15 source selections.

As a result of protest fear, technical evaluations appear to be somewhat compromised. This is important since compromised technical evaluations also decrease contractor performance. This construct assessed phenomena such as (1) technical evaluators not being allowed to say what needs to be said in a technical evaluation, (2) constraints imposed on the evaluations impeding the ability to write a meaningful evaluation, and (3) upon evaluation of proposals, a technical evaluator recognizing a need to change at least one evaluation criterion or its definition. Additional training for the technical evaluators could help increase their level of competence within the evaluation process. The evaluation process involves many people that are not necessarily familiar with the case law and pitfalls,
giving rise to bid protests. Additionally, the technical individuals that determine and define
the evaluation criteria should be the same individuals that evaluate proposals (i.e., apply the
criteria). Current, detailed, and standardized training for technical evaluators should result in
better-defined evaluation criteria and better application of them to proposals.

While a fear of protest did not affect perceived source selection method
appropriateness, protest fear was associated with the inappropriate use of the LPTA source
selection method. In turn, LPTA inappropriateness negatively affects contractor
performance. While these effects have been anecdotally espoused by practitioners, this
research is the first to quantitatively test the postulates. There were 23 respondents (7.5%) that
revealed that the source selection method used was to some degree inappropriate.
While this proportion appears small, it can be argued that any single instance of an
inappropriate source selection method gives room for pause. LPTA could be inappropriately
used since (1) evaluations can generally be accomplished more quickly and easily when
evaluated as pass/fail rather than by a subjective rating; (2) the government’s recent
increased focus on low price; and (3) the lower odds of receiving a bid protest compared to
arduous and mistake-prone procedures of a full tradeoff method. Further research should
confirm reasons why inappropriate source selection methods are employed, then acquisition
leaders should seek ways to mitigate those factors. Perhaps contracting officers should be
able to tap an independent panel of contracting professionals when they encounter leaders
or reviewers who will only approve a source selection method that does not correspond well
to the buying situation. This anonymous panel would then insert its documented opinion into
the contract file.

These findings are also germane to contractors. When a buying office concocts an
acquisition strategy that appears ill-suited to the buying situation (e.g., LPTA versus full
tradeoff for a highly complex requirement), it may be due to the fear of a bid protest.
Prospective offerors may misinterpret the use of LPTA as an added emphasis on price.
Their bid strategy, then, may be influenced by reducing costs and price, thereby putting high
performance at risk. Whereas, the agency may not actually be terribly concerned about
price.

The fear of protest diminishes a contracting officer’s perceived authority (i.e.,
discretion in making decisions). This is important since diminished contracting officer
perceived authority was found to decrease contractor performance directly. Contracting
officer’s perceived authority also affects contractor performance indirectly by decreasing
compromised technical evaluations. Many decisions and source selection documents
receive scrutiny via a litany of outside reviews (e.g., supervisors, peer review, contracts
committees, legal). Often, legal and committee advisors will conservatively require wording
changes to documents, changes to ratings, amendments to the request for proposals,
further discussions to clear up any uncertainty in evaluations, and retain offerors in the
competitive range—just to name a few. This level of oversight is another signal of the
importance the government places in avoiding a bid protest. Admittedly, it also coincides
with a less competent acquisition workforce (Punaro, 2012). Rather than treating the
problem, however, the symptoms gain the attention. Fixing the problems of contracting
officer competence and a cumbersome source selection process is difficult and lengthy.
Adding oversight is quick and simple. The implications are clear; better training is needed for
contracting personnel and technical evaluators to develop the requisite competence in
source selections, then oversight and reviews should be curtailed. Some protest risk must
be accepted for the sake of efficiency and better decision-making (i.e., negotiations and
award determination) leading to higher contractor performance.
This research confirmed the presence of outside influence on acquisition strategy decisions, and these influences carry associated implications for contracting. One interview informant commented, “I will tell you, legal pushes the LPTA. They push it a lot.” One survey respondent offered, “At this juncture, there are too many hands in the soup, and the procuring contracting officer (PCO) authority has been diminished. Attorneys need to resume the role of counselors again.” Since the source selection method is not a matter of legal sufficiency, attorney influence is curious. Selecting the source selection method is a contracting officer’s decision based on experience, knowledge, and professional judgment. Otherwise, government agencies may employ a costly professional contracting workforce with a high degree of accountability but diminished authority. If not capable, trusted, and empowered to make the necessary decisions, procurement clerks (e.g., Series 1105) would be much less costly than contracting professionals (e.g., Series 1102).

The fear of protest is positively related to an increase in transaction costs. Costs were assessed in terms of the number of personnel involved in a sources selection and their allocated time. The average cost per source selection was $235,236 (median = $165,832) with a standard deviation of $291,620. Notably, these costs are understated by considering direct salaries only; they exclude the true burdened cost of a government employee. An average of 9 different people worked on a given source selection team in the various roles (an average of 3.5 full-time equivalents). As a percentage of the total contract price, the transaction costs averaged 7.7% (median 1.2%). Compared to common interagency surcharges for contracting services (that cover post-award administration costs in addition to sourcing costs) of 1%–8%, these sourcing-only costs seem excessive. Thus, agencies may be operating at costs well above their collected fees, and these costs can be traced to fear of protests.

Post hoc analysis showed that as the fear of protest increases, the number of personnel and the actual procurement lead time increase. From the data, the average planned PALT was 183 days. The average actual PALT was 237 days. The difference, 54 days, constitutes added transaction costs. Thus, efficiency is compromised with greater fear of protest. While these salary costs may be dismissed as sunk costs, certainly excess personnel could accomplish other pertinent work if not serving on the source selection team for an extended time. These opportunity costs should not be ignored—particularly given the ubiquitous, persistent failures in other areas of acquisition such as contract administration (DoDIG, 2009b). If measures can be taken that reduce the fear of protest, transaction costs can be decreased. Likewise, if evaluation, negotiation, internal reviews, and documentation processes can be streamlined and if agencies can accept more protest risk, perhaps lead time can be saved, resulting in reduced transaction costs. Given today’s budget constraints and highly-leveraged financing, the significant transaction costs associated with source selections should not continue to be ignored. A first step would be to capture the quantified resources required to execute a source selection in a contract action reporting database (e.g., FPDS-NG). Agencies could also follow the for-profit sector’s lead by assessing and publishing key metrics such as total spend per sourcing full-time equivalent (CAPS Research, 2011).

These results surrounding transaction costs raise questions concerning the acquisition process in general. For instance, the single criterion for new case law—and hence, new reactive policies and regulations—is fairness, with no regard for efficiency. Is there a ceiling cost on fairness? Is there a point at which fairness is too costly? Additionally, the high amount of transaction costs suggests that the drivers of those costs be considered. Can policies, procedures, laws, case law, and regulations be reexamined and streamlined.
without compromising fairness? Is government procurement at the point of a source selection overhaul with a keen eye toward efficiency?

References


Federal Procurement Data System–Next Generation (FPDS-NG).


Avoiding Terminations, Single Offer Competition, and Costly Change Orders with Fixed-Price Contracts

Andrew Hunter—is a senior fellow in the International Security Program and director of the Defense-Industrial Initiatives Group at CSIS. From 2011 to 2014, he served as a senior executive in the Department of Defense, serving first as chief of staff to under secretaries of defense (AT&L) Ashton B. Carter and Frank Kendall, before directing the Joint Rapid Acquisition Cell. From 2005 to 2011, Hunter served as a professional staff member of the House Armed Services Committee. Hunter holds an MA degree in applied economics from the Johns Hopkins University and a BA in social studies from Harvard University. [ahunter@csis.org]

Gregory Sanders—is a fellow with the Defense-Industrial Initiatives Group at CSIS, where he manages a team that analyzes U.S. defense acquisition issues. Utilizing data visualization and other methods, his research focuses on extrapolating trends within government contracting. This requires innovative management of millions of unique data from a variety of databases, most notably the Federal Procurement Database System, and extensive cross-referencing of multiple budget data sources. Sanders holds an MA in international studies from the University of Denver and a BA in government and politics, as well as a BS in computer science, from the University of Maryland. [gsanders@csis.org]

Abstract

Fixed-price contracts offer the promise of controlling costs but are less likely to succeed when there is uncertainty regarding requirements. While these broad principles are uncontroversial, disagreement rages regarding the practical question of how widely they should be used. This study tests a variety of hypotheses regarding what contract characteristics are associated with better performance under fixed-price contracts. Here, performance is measured across three dependent variables: (a) the Number of Offers Received for competed contracts, (b) whether the contract was terminated, and (c) the extent to which change-orders raised the contracts’ cost ceiling. The study team has created a Bayesian network, populated by completed, publicly reported DoD contracts from FY2007 to FY2013 to address this research question.

The public purpose of this process also includes facilitating future acquisition research on a range of topics. To support future research, all analytical data and codes developed and/or used are posted on the CSISdefense GitHub (Sanders, 2015). This resource addresses two vexing issues that bedevil a wider use of the Federal Procurement Data System (FPDS) by academia, government, and industry researchers, namely (a) the data-selection barrier to using the FPDS and (b) the difficulty of deriving performance outputs from FPDS.

1 The Center for Strategic and International Studies (CSIS) does not take specific policy positions; accordingly, all views expressed in this presentation should be understood to be solely those of the author(s).
Introduction

The current preference within the executive branch and congressional oversight committees for fixed-price contracts is a reaction to external budget pressures and cost overruns associated with certain high profile acquisition programs (Younnossi et al., 2007). But even the DoD acknowledges that fixed-price contracting is not a panacea for all that ails federal acquisition. Discussions on the role of fixed-price contracting in acquisition is really a discourse on requirements, risk allocation, and uncertainty. Accordingly, as a contracting vehicle, fixed-price contracting will find more utility in specific, well defined scenarios such as in events where requirements and likely costs are well understood. More so, than in contracting situations characterized by uncertainty. This issue is acknowledged by OMB which has issued the guidance governing such acquisition transactions (Office of Management and Budget, 2009).

Under what circumstances are fixed-price contracts most likely to succeed? This paper addresses that question by examining seven years of DoD contract transactions available through the Federal Procurement Data System (FPDS). The project compares the performance results of fixed-price and cost-based contracts across a range of characteristics that typically would be known to the contracting officer before the contract was signed. While FPDS does not directly measure performance, it does capture three variables that directly measure the possible drawbacks of fixed-price contracts. These are

- Single Offer Competition,
- Contract Terminations, and
- Cost-Ceiling-Raising Change Orders.

“Single-Offer Competition” can be indicative that the pricing vehicle involves too much risk on the vendor side, which makes bidding on the contract unattractive to potential vendors. Tracking “Contract Terminations” addresses the risk of outright failure because a

---

2 Cost growth in defense acquisition is driven by several factors that include schedule, unrealistic estimates, acquisition strategies and funding availability. Studies of weapon system cost growth, notably by researchers indicate that across board the average adjusted total cost growth for a completed program was 46% over the study period of 30 years.

3 On numerous occasions, Frank Kendall, the under secretary of defense for Acquisition, Technology, and Logistics, has underscored the need to use appropriate contracting vehicles by pointing out the problems that a fixed-price contract caused during its development of the A-12 Avenger.

4 The concept “fixed price” deserves some further illustration. A fixed-price contract suggests a price that is not subject to any adjustment on the basis of the contractor’s cost experience in performing the contract. And according to provisions in the Federal Acquisition Regulation (FAR), a firm-fixed-price contract is suitable for acquiring commercial items (see Parts 2 and 12) or for acquiring other supplies or services on the basis of reasonably definite functional or detailed specifications (see Part 11) when the contracting officer can establish fair and reasonable prices at the outset. But as is succinctly noted in the DoD’s (2014) Performance of the Defense Acquisition System: 2014 Annual Report (pp. iii–vi), “Prices on fixed-price contracts are only ‘fixed’ if the contractual work content and deliverables remain fixed”; such contracts can be (and often are) easily modified to handle unexpected technology gaps, engineering issues, or shifting threats, leading to cost growth. At times fixed-price vehicles can be virtually indistinguishable from cost plus vehicles, as was the case with the Air Force’s canceled Expedtionary Combat Support System (ECSS). This reading of “fixed price contract” by the DoD guides our understanding of fixed-price contracts in the framework of this study.
vendor has taken on more risk than they can handle. And finally, “Cost-Ceiling-Raising Change Orders” demonstrates that the government customer finds the present contract structure unsatisfactory and may indicate cost overruns.

The study team analyzes the drivers of these three dependent variables using a Bayesian network populated with all unclassified\(^5\) completed DoD contracts from FY2007 to FY2013. The study objective is to determine when fixed-price contracts are most effective. The team drew hypotheses from the literature regarding under what circumstances the fixed-price mechanism would lead to better results. The team also developed control variables. These controls divide the data to capture factors not relating to pricing mechanism that may lead to better or worse performance.

Addressing the question of when fixed-price contracts are most effective is only a portion of the public purpose of this project. The data that undergird this investigation originate from FPDS and are thus open source and pertinent to a wide range of government contracting questions. With this chosen approach, the study team simultaneously addresses two of the most vexing issues that have bedeviled the wider and more effective use of FPDS by academic, government, and industry researchers: the high barrier to entry for accessing complete and relevant data and the difficulty of deriving performance outputs from FPDS.\(^6\) To ensure reproducibility of this analysis and to provide a starting point for future research, the entirety of the dataset is published through the CSISdefense GitHub account along with the typographies and analytical code that the study team used to create the statistical models (Sanders, 2015).

Background

Since the inception of the federal acquisition business, issues of performance have vexed policy-makers. In the last years, major data analytics work by DoD analysts found scant statistical correlation between industry profit margins and program performance in cost and schedule. For example, rigorous analysis of data from hundreds of major weapons programs strongly suggest that hitherto fixed-price contracting does not always achieve the set goals (DoD, 2014). As far back as 1949, John Perry Miller, in Pricing of Military Procurements, exposed the lack of symmetry between incentives and contract performance (as cited in Williamson, 1967, p. 218). In the same hue, scholarly works from the early 1960s by Frederick M. Scherer (1964) and Peck also support the notion that there is a discrepancy between values attached to incentives and performance results (Kaysen, 1963). Speaking to this issue, Williamson (1967) glumly stated,

> My analysis of these relations leads me to conclude that neither the manipulation of profit incentives, nor the monitoring of contract progress can be expected, in any dependable sense, to yield significant improvements in contract performance as long as the specification of the task remains unchanged. (pp. 217–218)

Procurement has long been at the top of the government’s laundry list of activities that need improvement. However, this policy focus has come in cycles—with attendant ups

\(^5\) There is no regulatory mandate to report classified contracts; as a result, CSIS assumes that they are not included within FPDS.

\(^6\) For a more detailed remark on these factors, please see the following section of this report.
and downs over the years. Practically every administration since the 1970s has embarked on some form of acquisition reform or at least paid lip-service to the process. Over 150 major studies devoted to the field of acquisition reform have been produced since the end of the Second World War (Schwartz, 2014). Despite these efforts, according to Congressional Research Service analyst Moshe Schwartz, DoD development contracts since 1993 have experienced a median of 32% cost growth—not adjusted for inflation. Since 1997, 31% of major defense acquisition programs have had cost growth of at least 15%. Schwartz observed that every year between 1996 and 2010, the Army spent more than $1 billion on programs that were ultimately canceled.

In 2009 and 2010, these problems led to the latest round of reforms, including the Weapon System Acquisition Reform Act and Better Buying Power. The first set of DoD Better Buying Power initiatives called for greater use of fixed-price incentive fee contracts when appropriate (Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics [OUSD(AT&L)], 2010). The second iteration of these reforms in 2012 suggested fixed price for low-rate initial production but emphasized contracting officer discretion in choosing the right pricing mechanism (OUSD[AT&L], 2012). In 2013, the annual report on the Performance of the Defense Acquisition System reinforced this change, finding that use of fixed-price contract types were not themselves correlated with cost savings. Throughout this debate, all sides agree that fixed-price contracts are more appropriate in some cases than in others.

Study Dataset and Performance Measures

The dataset for this study consists of DoD contracts reported in FPDS that were initially signed no earlier than FY2007 and completed by FY2013. Notable exclusions include classified contracts not mandated to be reported in FPDS, contracts funded but not managed by the DoD, and Defense Commissary Agency contracts which have not been reported in recent years. To enable comparisons, the dataset is not limited to fixed-price contracts.

Determining when contracts are completed is the most challenging portion of compiling the dataset. Contracts closed out or terminated by the end of FY2013 are included even if their current completion dates run into the next fiscal year. However, many contracts in FPDS and in the sample are never marked as closed out or terminated in the Reason for Modification field. In these cases, completion status is based on the current completion date of the most recent transaction in FPDS. This method could accidentally include contracts that have not reached their ultimate conclusion dates and are merely dormant. However, the FY2013 sample end date means that any such contracts would have to be inactive for an entire fiscal year, which is remarkably unlikely.

FPDS raw data is available in bulk from USAspending.gov starting in FY2000. However, data quality steadily improves over that decade, particularly in the commonly referenced fields of interest to this study. In most cases, unlabeled rates topped out at 5% to 10%. The critical exceptions are the Base and All Options and Base and Exercised Options fields, which report contract ceilings. Prior to FY2007, these fields are blank for the majority

7 These exclusions are common with other DIIG work. See DIIG methodology for more details: http://csis.org/program/methodology
of contracts. When that field is not available, calculating the percentage growth rate due to cost-ceiling–raising change orders is impossible. In addition, this study classifies contract size by initial ceiling and not total obligations because the latter figure is dependent on contract performance.

Because a key dependent and independent variable are not available prior to FY2007, the study team chose to set FY2007 as the start date rather than risk sample bias by including only those earlier contracts which were properly labeled. This restriction poses a significant limitation in that no contracts of more than seven years in duration can be included and five-year contracts are only in the study period if they started by October 1, 2007, or were closed out early.

The dependent variables, as well as the other study variables, are available through the project's GitHub page (Sanders, 2015). The entire dataset includes nearly six million contracts, and as a result the study team will also provide smaller samples as a more accessible starting point. The means of sampling will be determined by consultation with members of the likely user community at the NPS conference and the interviews during the refinement phase of this project.

The largest sampling challenge is that approximately two-thirds of the contracts have a total obligated value less than $25,000, but these contracts represent only about 3% of the obligations for the period. As a result, when this paper deals with samples for computation reasons, the sample is weighted by total obligations to better reflect DoD spending patterns rather than merely describing the numerous, but comparatively insignificant in aggregate, small contracts. The Bayesian model developed later in this paper is calculated using the complete dataset but includes a variable differentiating by initial contract ceiling, which is intentionally constructed to ease study of larger contracts.

This approach addresses two of the largest obstacles to wider use of FPDS within the government, academia, and industry. First, the two official portals, USAspending.gov and the FPDS web-tool, both perform a critical service in giving access to contract transaction in aggregate or detailed form. However, for many researchers the relevant unit of analysis is contracts and not transactions. Both websites can be used to access the full records of individual contracts, but due to data inconsistencies and bulk download restrictions, they are not well suited to larger sample studies.

The limitations of the data in raw form can be overcome by downloading the complete data feeds via the data tab of USAspending.gov, but with each year accounting for multiple gigabytes, this represents a high barrier to entry for researchers who lack the tools or training for large dataset work. This challenge is further increased by the often arcane nature of the data fields and the need to undertake cleanup and refer to multiple columns to get to data of interest. This first challenge is attested to by the regular calls the study team receives from other researchers seeking to use FPDS.

The second obstacle is that FPDS almost exclusively measures contract inputs but not performance outputs. Measures of contract performance do exist in other databases, but they are largely inaccessible without, at very least, an official government purpose or permit. This project takes a step towards overcoming that problem through three dependent variables referenced earlier in this section. These variables were chosen due to their relevance to fixed-price contracting and availability, but they are also applicable to a wider range of research questions. Due to their importance and broader reference, this paper discusses each of the three in detail in a subsequent section.
Single-Offer Competition

Whether or not a contract is competed is primarily an input into performance. However, single-offer competition also reveals information about the request for proposals. A solicitation that only has a single respondent indicates some combination of three factors: thinness in the underlying market; a failure to notify or give adequate response time to potential competitors; or a contract that is unappealing to vendors. That final point is of interest to this study. Fixed-price contracts transfer risk to vendors, and if vendors perceive greater risk than the government is willing to pay for, then fewer vendors may be willing to bid.

Methodology

Non-competed contracts are not included in this analysis because the choice of whether or not to compete is based on factors that are already known before the choice between fixed-price and cost-based contract is made. For example, when a sole-source award is justified based on there being “only one source,” that rationale refers to the total number of potential vendors and not the number of interested vendors. As a result, for the graphs on single-offer competition and for analysis at the end of this piece, sole-source awards are excluded. While multiple variables are used to judge whether a contract has been completed, only one, the Number of Offers Received, is necessary to determine how many vendors really did submit an offer.

Whether a contract is competed is calculated largely using the same approach as in prior CSIS studies ("Methodology," n.d.). This method emulates the official DoD methodology to the extent possible when using raw data downloads rather than the FPDS webtool. In the vast majority of cases, competitive status is classified for the entirety of the contract duration. Thus if a contract had a duration of three years and was competed in the first year, it qualifies as competed for the entirety of the duration. This also extends to single-award indefinite delivery contracts, which are classified based on whether the original vehicle was competed rather than consistently treated as only receiving an offer from the single awardee. The Number of Offers Received is calculated using the same strategy.

To see more details on the construction of the number of offers, visit Contract_Competition.md on the project’s CSISdefense/Fixed-Price GitHub page (“Fixed-price/Contract_Competition.md,” 2015). Summary charts are included in the parallel CSIS report on Competition, which is also in the conference proceedings.

Terminations

Abruptly ending a contract through termination is a challenging endeavor for the government. The proximate cause of the termination may not be vendor performance, but instead a drastic change in government needs, the failure of a related contract, or the cancellation of the entire program. However, in all three cases the government has the option of simply running out currently exercised options and stopping further payments. Thus, even if the source of the failure was outside the contract, a termination indicates that the contract was unable to adapt to changing circumstances. Critically, the greatest vulnerability of fixed-price contracting will result in a termination: if too much risk is placed on the vendor and they outright fail beyond the point at which adjusting the contract could turn things around. In this instance, the government may lose any resources that have already been invested as well as paying a significant cost in time to start the project over.
Methodology

Contract termination is determined using the *Reason for Modification* field in FPDS. A contract is considered to be terminated if it has at least one modification with the following values:

- “Terminate for Default (complete or partial)”
- “Terminate for Cause”
- “Terminate for Convenience (complete or partial)”
- “Legal Contract Cancellation”

These four categories and the “Close Out” category are used to mark a contract as closed. As is discussed above, many contracts well past their current completion date never have a transaction marking them closed, however, a termination is an active measure that mandates reporting unlike the natural end of a contract which can go unremarked.

The four different values of contract termination provide useful granularity, but for aforementioned reasons even a partial termination for convenience indicates that something has likely gone awry. Thus, given the already low number of terminations, the study team treats a contract as either terminated or it is not, rather than subdividing by type.

To see more details on the construction of the number of offers, visit Contract_Competition.md on the project’s CSISdefense/Fixed-Price GitHub page (“Fixed-price/Contract_Competition.md,” 2015).

Cost-Ceiling-Raising Change Orders

Change orders are not as severe an indicator of trouble as terminations. A change order might result from a contract being adapted to a changing environment or even being adapted to further take advantage of a successful innovation. Even when the change order indicates a mistake, it may often not be on the vendor side. Instead, requirements creep prompted by the government may add expensive new tasks to the contract. The affordability of fixed-price contracting comes in part from their simplicity and inflexibility. Thus, when fixed-price contracts are subjected to a large number of change orders, whether prompted by government or vendor actions, this is a warning sign that a different form of pricing may have been more affordable.

Methodology

Similar to contract terminations, change orders are reported in the *Reason for Modification* field. There are two values that this study counts as change orders: “Change Order” and “Definitize Change Order.” For the remainder of this report, contracts with at least one change order are called Changed Contracts.

There are also multiple modifications captured in FPDS that this current study will not investigate as change orders. These include:

- Additional Work (new agreement, FAR Part 6 applies)
- Supplemental Agreement for work within scope
- Exercise an Option
- Definitize Letter Contract

The Number of Change Orders refers to the number of FPDS transactions for a given contract that lists one of the two change order categories as their *Reason for Modification*. The vast majority of contracts do not receive change orders, but changed contracts are still far more common than terminations.
This study uses changes in the *Base and All Options Value Amount* as a way of tracking the potential cost of change orders. The *Base and All Options Value Amount* refers to the ceiling of contract costs if all available options were exercised. The alternative ceiling measure, *Base and Exercised Value Amount*, is not used because contracts are often specified such that the bulk of the eventually executed contract, in dollar terms are treated as options. In these cases, the all-inclusive value provides a better baseline for tracking growth.

The *Obligated Amount* refers to the actual amount paid to vendors. This study team does not use this value for this analysis because spending for change orders is not necessarily front-loaded. For example, a change to a contract in May of 2010 could easily result in payments from May 2010 through August 2013.

The % Growth in Base and All Options Value Amount from Change Orders is calculated as follows:

\[
\text{% Growth} = \frac{\text{Base & All Options Value Increases from Change Order Modifications}}{\text{Base & All Options Value Amount for Original, Unmodified Transaction}}
\]

To see more details on the construction of the number of offers, visit Contract_ChangeOrders.md on the project’s CSISdefenses/Fixed-Price GitHub page (“Fixed-price/Contract_Competition.md,” 2015).

**Bayesian Network Model Building**

A variety of statistical techniques are appropriate for inferential analysis on when fixed-price contracts are most likely to be successful. The study team chose a Bayesian network approach for three key reasons. First, this approach scales well to large datasets, such as the nearly six million defense contracts completed between FY2007 and FY2013. Traditionally, only a sample of such data would be available, but thanks to FPDS and modern computing, it is possible to analyze the entire population. Second, while a Bayesian network approach and other similar techniques can be used for prediction, it is particularly well suited to understanding how the different pieces of evidence are interrelated. Because this project seeks to provide a starting point for future research, enhancing understanding of the model’s causal logic is more important than creating a model which optimizes the ability to predict outcomes.

Finally, the knowledge engineering process used with Bayesian models—building connections between evidence, called whitelists and blacklists, and the subsequent model queries—is well suited to CSIS’s strength in accessing acquisition domain experts and data scientists. The initial model in this paper will be built upon over time in two ways. First, it will be expanded from examining only the Number of Offers Received to including all three dependent variables. Second, it will be refined by consultation with additional external experts and listening to their insight about which pieces of evidence should be connected and where more granularity may be appropriate.

The model is built in the open source statistical programming language R using two modules. The module BnLearn is used for the Bayesian network learning process, which turns the collected data into a directed graph that is acyclic, which is to say there are no loops (Scutari, 2010). The module gRain is used for the second part of the process, creating the conditional probability table and then querying the resulting multiples (Højsgaard, 2014). Both modules are also open source and the data as well as the processing and analytic programming code used to implement this process are available through the CSISdefense fixed price GitHub repository (Sanders, 2015).
**Description of Evidence**

As with most statistical models, the first step with a Bayesian network is to gather, clean, and transform the data. Each piece of evidence was collected by first applying CSIS transaction-level lookup tables. Once the transaction level data was categorized in SQL server, it was then collated into contract-wide values. The first step in this process was done using codebooks available at the lookup-tables repository of the CSISdefense GitHub account. The combination stage was done on a variable by variable basis and that process is covered in the fixed-price repository of the CSISdefense GitHub account. The last stage was conversion into evidence notes, each having between two and eight distinct states. The processing required for a Bayesian network increases exponentially with each new node of evidence. As a result, the initial model is intentionally minimalist, and our study team will add more granularity to the model as developments warrant.

**Contract Fundamentals**

These nodes of evidence are largely set by the needs of the relevant portion of the Department of Defense rather than chosen by the acquisition official. There are choices to be made within them, for example, whether to fill a need directly with as a product or via a service provider. However, as a rule these nodes of evidence influence the contracting method and not vice versa.

**Who (Component: Army/Navy/Air Force/Other DoD):** Determined by the contracting office rather than the funding office. This will be referred to as “component” throughout this discussion.

**What (Platform: Air/Land/Vessel/Electronics & Communications/Missiles and Space/Weapons and Ammunition/Facility Related Services & Construction/Other):** Determined by the combination of the claimant program code for the platform when available and otherwise via the product or service code. This will be referred to as “platform” throughout this discussion.

**PSR (Product/Service/R&D):** Determined by the product or service codes, with R&D management and support being treated as a service.

**Intl (International: Just U.S./Some International):** Based on the place of performance. Those contracts with any transactions in foreign countries are treated as having some international. Lookup tables are primarily used when imputing data is necessary due to a missing or malformed value in either the country or U.S. state place of performance fields.

**Link (# of linked contracts: none/1–749/750+):** This calculated column is the study team’s first effort to account for the possibility of problems cascading from other related contracts. For those contracts without a system code, the value is set equal to the number of preexisting contracts in the same contracting office that share a Platform characterization (excluding the Facility-related Services and Construction [FRS&C] and other platforms). For those contracts with a system equipment code, the value is equal to the number of preexisting contracts sharing the system equipment code. This value is then supplemented with the number of contracts with the same platform and contracting office (with the aforementioned exclusions) that are not labeled with any system or equipment code. This field will be referred to as “Interlinkages” throughout this discussion.

See the fixed-price repository of the CSISdefense GitHub account for processing code used for this calculation.
**Contract Approach**

The contract approach refers to those contract characteristics that are chosen by the relevant acquisition officials in the pursuit of a successful outcome.

**Comp (Competition: Comp./No Comp.):** Is determined using the standard CSIS methodology, with the critical exception that the numbers of offers received is treated as a separate piece of evidence.

**Ceil (Ceiling: $15,000/$100,000/$1,000,000/$30,000,000):** Refers to the initial ceiling on total potential contract obligations. Is set by the initial Base and All Options Value for the contract. This value was chosen rather than the initial Base and Exercised Options Value because exercising options happens regularly during the course of an on-time and on-budget contract.

**Dur (Duration: One day to Two Months/Seven Months to a Year/More than a Year):** Refers to the duration and is calculated using the number of days between the initial effective date and the current completion date for the contract. The ultimate completion date is also available but was regularly unlabeled.

**FxCb (Fixed price or Cost-base):** Fixed price includes all forms of fixed-price contracting except fixed-price level of effort. That comparatively rare form has been described in meetings with DoD officials as exhibiting more properties of cost-based contracts. Cost-base includes all forms of cost-plus contracts as well as time and materials and labor-hours contracts.

**Indefinite Delivery Vehicle (IDV):** Indicates whether or not a contract is one of the many forms of Indefinite Delivery Vehicles (IDVs). This is a contracting approach in which a single root contract is used as a basis for multiple other contracts.

**Contract Outcomes**

**Offr (Number of Offers Received: 1, 2, 3–4, 5+):** This is the first dependent variable and is described in greater detail earlier in this document. Cases with no competition are categorized as only one offer, although for hypothesis purposes, the existence of competition will typically be included in the model query. The remaining dependent variables will be added in to the model in coming research stages.

**The Whitelist of Mandatory Arcs Between Pieces of Evidence**

The whitelist is a collection of directed arcs between pieces of evidence that must be included in the final model whether or not the learning algorithm recommends them. Developing the whitelist was an iterative process. First the team determined which pieces of evidence were most strongly linked and then, after seeing initial results, added further connections where the learning algorithm could not determine the flow of causality. Finally, the team compared different versions of the model created using variant algorithms and added those arcs which the study team deemed important but that were absent in some models. The whitelist in Figure 1 is the final one at the time this paper was written, including all three iterations of building the list.
In interpreting this graph, each circle is a piece of evidence or node. Each arrow is a directed arc, and the parent node influences the child node that the arrow points to. For example, there is an arc from the parent Component ("Who") node to the child Platform ("What") node because the different components buy different mixes of platforms. Thus if Component is “Army” then the Platform is substantially more likely to be “Land Vehicles” and less likely to be “Vessels.” So long as no loops are formed, each node can be linked to multiple or no other nodes, for example Product/Service/R&D (PSR) influences both “Fixed price or Cost-base” and Interlinkages (“Link”) and “Number of Offers Received” (Offr) has arcs coming from both comp and “Fixed price or Cost-base.”

Working from the bottom, “Number of Offers Received” has “Fixed price or Cost-base” and Comp as its parents because the number of offers directly depend on whether a contract is competed and “Fixed price or Cost-base” is the study variable, and thus the connection with the dependent variable is highly of interest. On the next level up, Interlinkages is a constructed variable and both Platform and PSR are used to create it. Who also plays a role, but primarily at the contracting office level and thus the connection is not as direct. "Fixed price or Cost-base’s” arc from PSR was there from the start, as R&D contracts are classically the domain of cost-based pricing. "Fixed price or Cost-base’s” second parent, Ceiling, was added near the end of the process because “Fixed price or Cost-base” only had one parent in some of the models and consistently including Ceiling in the rest. Further analysis via cross-sectional graphs showed that after accounting for PSR, Ceiling appeared to have the strongest influence on “Fixed price or Cost-base.” Namely, contracts with high ceilings are notably more likely to be Cost-Based or Combination.

Platform is influenced by Components in another straight forward connection as was covered in an earlier example. IDV is a parent of Ceiling because the learning algorithm could not decide on the direction of the relationship. The study team decided that IDV was the parent because choosing an IDV can often mean choosing to break a goal into multiple smaller pieces, each with a small scope. Thus the Ceiling of a project can depend on whether or not the contracting officer feels an IDV is available and appropriate. The last two nodes, Intl and Duration, do not have any whitelist entries at this time. During model creation, the study team experimented with linking Duration to “Number of Offers Received,”
but that caused the learning algorithm to reject the stronger connection between Ceiling and “Number of Offers Received.” The team could force both connections, but it is not necessary to do so and the relationship between Duration and “Number of Offers Received” appeared complex in cross-sectional graphs. As will be shown later, this does leave Duration without any children, but this problem will likely be remedied once additional dependent variables are added.

**The Blacklist of Mandatory Arcs Between Pieces of Evidence**

Black lists are the inversion of white lists: arcs that may never be included in the model regardless of the findings of the learning algorithm. Developing the blacklist was similarly an iterative process, although with one notable exception, most of the revisions were merely adding more arcs to the blacklist to correct for possibilities overlooked in prior iterations.

For Figure 2, the red lines indicate arcs that are not allowed. For example the arrow from “Number of Offers Received” to Platform means that the number of offers cannot have a causal influence on what sort of platform is being bought. In many cases, this is for straightforward causal reasons. The evidence regarding contract fundamental can influence the contracting approach, but not vice versa. This graph is has many more arcs than the prior Whitelist graph because it is straightforward for experts to establish which evidence factors are decided earlier along the timeline or are take precedence over deciding other related factors. For example, the relevant acquisition official will typically first determine whether they can compete a given contract and only then determine what vehicle or pricing mechanism would be appropriate.

The existence of a blacklist arrow does not mandate that there is a connection going the other direction. In fact, that is the point of the blacklist, to prevent spurious connections from being made without committing to an arc going in the opposite direction. For example “Number of Offers Received” is blacklisted to every other piece of evidence in the model, because competition takes place only after the other factors are set in broad terms. However, as will be seen in the causal model, not every piece of evidence that can arc to “Number of Offers Received” does so.
In a smaller number of cases, the blacklist extends to arcs in both directions. This means that the two evidence nodes, connected with a purple arc and an arrow pointing in both directions cannot be parents nor children of each other. In the case of IDV and linked, this is because the causal link is difficult to adjudicate because of confounding factors. For contracting office and platform pairings that have IDVs available, the number of preexisting interlinked contracts will typically be higher. However, that reflects the omitted variable of whether there are active IDVs rather than a direct connection between the two evidence nodes.

Component evidence node has bidirectional blacklisted arcs with all of the contracting approach evidence nodes as well as with the “Number of Offers Received” outcome node. The study team chose to block these links because the literature review did not find a theoretical basis for the organization itself, rather than the characteristics of its contracts, being a key determiner for fixed-price success or failure. CSIS is separately examining that question in its report on DoD components, which is also being presented at this conference. Pending the outcome of that study, these bidirectional blacklisted arcs may be revisited.

Initial Results

After the iterated whitelist and blacklist generation process, Figure 3 shows the ultimate result. In the figure, the blue arrows are those arcs that were mandated by whitelists. With the exception of links to Duration from Ceiling, IDV, and Comp, the direction of the remainder of the arcs was locked in by the blacklist. Thus, PSR did not have to have a direct connection to Duration, but the alternative was disallowed.

The resulting Bayesian network is highly interlinked, as is shown by the sheer number of arcs, well in excess of the small number required by the whitelist. However, there
are two notable areas where connections are sparse. Duration is influenced by six different evidence nodes but does not influence Offr or any other node. This may change in the future as a result of introducing greater granularity to Duration or after the introduction of the other dependent variables. The second evidence node that is remarkably isolated is "Fixed price or Cost-base," which suggests humility may be necessary regarding the influence of the study variable.

![Initial Number of Offers Bayesian Network](image)

**Figure 3. Initial Bayesian Network**

**Preliminary Hypothesis Testing**

This initial model allows for preliminary testing of four of the five study hypotheses as well as robustness checking using five different controls. The study team developed each of these hypotheses and controls from the literature, rather than through learning algorithm described above.

As presently configured, the various evidence nodes are not granular enough to fully test the hypotheses proposed in the early stages of this research. However, four of the five hypotheses can be tested. These hypotheses were often formulated with a greater level of specificity than the model presently allows, for example, examining contracts with a ceiling of greater than $500 million or addressing only software contracts. Categories which only apply to a small fraction of contracts can increase the complexity of the model while reducing its statistical power. This trade-off can be worthwhile, but it will first be tested with model variants and only when most insightful will these additional breakdowns be included in the model as a whole.
The current Bayesian model enabled five different controls which were employed by comparing fixed price and cost-based Number of Offers Received for a different subsets of the contract dataset. As the Bayesian model becomes more inclusive, more controls, such as separating undefinitized contract actions, will be added.

- Past research has found aircraft and drone contracts to be especially challenging,\(^8,9\) so the first control split the dataset using the Platform evidence node (DoD, 2015; Ritchie, 1997).
- Both larger and smaller contracts sometimes show different trends than those in the middle. The second control separated out large contracts as those with a ceiling of/over $30 million. The third control separates out contracts with a ceiling of less than $1 million.\(^10\)
- Indefinite Delivery Vehicles are associated with higher rates of competition and are separated out for the fourth control.\(^11\)
- Finally contracts with a duration of greater than a year may have higher risk thus experience different dynamics.

**Hypothesis 1: Large R&D contracts will perform better as cost-based contracts.**

Our interim findings generally support both the academic literature\(^12\) and policy documents\(^13\) that posit large R&D contracts may perform better under cost-based contracts. As expected, cost-based contracts had a substantially lower single offer competition rate, 22.5% for cost-based versus 29.7% for fixed price. Cost-based similarly had a higher rate of competition with five or more offer. This pattern held for long duration contracts and for aircraft contracts. In the latter case fixed-price contracts received only a single offer 37.6%
of the time versus 11.3% for cost-based! In these categories and in the overall sample, combination contracts more closely resembled fixed-price contracts when it came to offers received. Cost-plus only loses its advantage for IDV contracts, where 26% of contracts receive only a single offer compared to 24.9% of combination contracts and 20.5% of fixed-price contracts. This discrepancy merits further study to determine what sort of IDV drives up the rate of single offer competition for cost-plus contracts.

Hypothesis 2: Complex projects, as measured by pre-Milestone B major defense acquisition project status, will perform better as cost-based contracts.

Complex projects, such as the development stage of Major Defense Acquisition Projects, are favored for cost based contracts in the literature (Bajari & Tadelis, 2001) but not supported by our analysis when the placeholder of linkages was used. However, the results were not highly robust and were contradicted for two of the five controls. The single offer competition rate for cost-plus contracts was six percentage points lower for large contracts and two percentage points lower for aircraft. Since both large contracts and aircraft and drone contracts are associated with MDAPs, this suggests that the study team will need to refine linkages or test the hypothesis directly by looking at system equipment codes.

Hypothesis 3: Contracts with a longer duration will perform better as cost-based contracts.

While prolonged schedules and unforeseen cost growth are reasons a vendor may prefer cost-based contracting, the hypothesized relationship did not hold. Unlike Hypothesis 2, the results appeared to be fairly robust. Overall, 25.3% of fixed-price contracts with durations greater than a year receive only one offer compared to 31.4% of cost-plus contracts. This gap narrows to less than a percentage point for large contracts and aircraft and drone contracts,

Combination contract rates resembled fixed-price rates in most categories, with the exception of aircraft and drones where more than half of all competed contracts received only one offer. Further analysis may be merited to determine whether the hypothesis holds when the threshold is set at two or more years.

Hypothesis 4: The potential for greater competition improves fixed-price performance.

The hypothesis that fixed-price contests are preferred by acquisition officials when they are likely to receive more competition may indicate that contracts that are more likely to be competed will perform better. This hypothesis is not testable with this dependent variable. However, given the high variability for the Number of Offers Received for fixed-price contracts and the number of evidence nodes influencing both competition and number of offers, this hypothesis should be straightforward to test in future stages.

14 Cost-plus contracts are preferred to fixed-price contracts when a project is more complex.
15 "But price redetermination might be used whenever contingency charges otherwise would be included in a contract price due to such factors as prolonged delivery schedules, unstable market conditions for material or labor, or uncertainty as to cost of performance" (Fixed Prices and Price Redetermination in Defense Contracts).
16 "[Government principals] would prefer a fixed-price contract when the number of bidders increases" (Goel, 2001).
Hypothesis 5: Large software projects perform better as fixed-price contracts.

Our results surrounding this hypothesis were inconclusive. We based the hypothesis on one piece of literature which stated that vendors prefer fixed-price contracts for software, contradicting other literature that hypothesized that vendors prefer cost-plus contracts for larger, more complicated projects. For this iteration, Electronics and Communications services were used as a proxy for software. We found that for small contracts, the hypothesis holds, with fixed-price contracts nearly 10 percentage points less likely to receive one offer compared to cost-based (28.8% versus 38.7%). However, this relationship reverses itself for large contracts; 29% of fixed-price contracts receive only a single offer versus 21.2% of cost-based contracts.

Conclusions

This promising hypothesis testing shows the value of the Bayesian network built by the study team. The analysis of single-offer competition for larger R&D contracts confirmed the conventional wisdom that vendors would be less likely to bid on fixed-price contracts but also revealed an intriguing wrinkle that this does not hold for large IDV R&D contracts. In addition, both the model building and the hypothesis testing surprisingly failed to uncover a relationship between contract duration and the number of offers a contract received. Both the literature and practitioners have often mentioned the importance of contract duration, and it may prove more consequential for contract terminations and ceiling-raising change orders which the study team will explore as this work continues.

Finally, the study team is excited to pursue the next steps of refining the model via external consultation and also expanding the model to include the remaining two dependent variables as well as a small number of intervening variables. The latest version of the model will be available at the CSISdefense GitHub account, and CSIS welcomes feedback from readers.18

References


17 “We hypothesize that the vendor’s ability to leverage information asymmetry about capabilities and experiences translates into the vendor preferring Fixed-Price contract to secure larger information rents. Our results support this hypothesis and suggest that the vendor would prefer the FP contract for larger and longer projects with larger teams. However, vendors would prefer a [Time and Materials] contract when the risk of employee attrition from the project team is high” (Gopal & Sivaramakrishnan, 2006).

18 Please visit https://github.com/CSISdefense/Fixed-price to see the latest version of the model and submit any questions or suggestions to gsanders@csis.org. For timing reasons any feedback that is not received by mid-summer 2015 is unlikely to be included in the final technical paper, but may be incorporated in future iterations of this work.


**Acknowledgments**

The study team would like to acknowledge the contributions of Jacob Bell, Gabriel Coll, Samantha Cohen, Ryan Crotty, Jesse Ellman, Guy Nzeribe, Madison Riley, and John Vick to the quality and robustness of this research effort.
Price Analysis on Commercial Item Purchases Within the Department of the Navy

Janie L. Maddox, Maj, USAF (Ret.)—has 30 years of experience working in the government contract management and educational fields. In 1994, she won the National Contract Management Association’s (NCMA) prestigious Blanche Witte Award for her leadership in the development of the Contract Pricing Reference Guides referenced in the FAR as the guidance for government contract pricing. Maddox is currently a professor at the prestigious Naval Postgraduate School. She lectures on federal government contract subjects, such as source selection and cost/price analysis. Her government teaching and educational management experience is mixed between both the private and government sectors. [jlmaddox@nps.edu]

LCDR Paul Fox—is a supply corps officer; He served aboard the USS George Washington (CVN 73) as the sales officer and materials officer. Follow on tours included food service officer for Joint Task Force Guantamano Bay, administrative contracting officer for the Defense Contract Management Agency Atlanta, and Defense Contract Management Agency–Iraq, Procurement Contracting Officer–Naval Regional Contracting Office–Naples Italy. [wpfox@nps.edu]

Ralucca Gera—is an associate professor of mathematics in the Applied Mathematics Department at the Naval Postgraduate School (NPS). She has been teaching and working on research at NPS since July 2005, during which she has advised 20 students, seven of whom worked on acquisition and contracting-related projects. Dr. Gera’s research is in networks, publishing 32 papers in reputable journals in graph theory and two more papers in the proceedings of the Acquisition Research Symposium. She was awarded department- and school-level teaching awards at NPS. Her research grants serve collaborations with the Computer Science Department and Graduate School of Business and Public Policy at NPS. [rgera@nps.edu]

Abstract

The objective of this project is to research current pricing memorandums and determine whether the use of price analysis techniques and pricing memorandums can improve acquisition-pricing outcomes. The purpose of the research is to explore the efficacy of the government’s current documentation of price analysis information. The intent is to diagnose the price analysis techniques that are being utilized and documented in the contracting file, and to explore potential improvements. This analysis is based on a review of a sample size of contract files and a personnel survey at a Department of Defense contracting office.

Introduction

The Federal Acquisition Streamlining Act of 1994 significantly changed government procurement practices. This act expanded the definitions of a commercial item and a non-developmental item, which in turn allowed vendors to avoid submission of certified cost and pricing data for commercial items in response to government contracting solicitations (Rumbaugh, 2010). One impact of this change is that contracting professionals must now look at market forces to establish price reasonableness for commercial item procurements. The importance of market research and price analysis methods has increased because of this change (Gera & Maddox, 2013). Since this change was made, the Department of Defense Inspector General (DoDIG) has issued multiple reports that are critical of contracting officers. Contracting officers have consistently failed to adequately justify price reasonableness, or they have failed to provide documentation that explains their price reasonableness determination. The purpose of this research is to explore which price analysis techniques are being utilized and documented in the contracting file, and to explore potential improvements within Department of Defense (DoD) Navy contracts. This project is a continuation of research accomplished on DoD contracts (Redfern, Nelson, & White, 2013; Gera & Maddox, 2013).
The objective of this project is to determine whether better use of price analysis techniques can improve DoD contract pricing. The intent is to diagnose both strengths and weaknesses and to explore potential improvements utilizing a contract file review and a personnel survey.

Many government audit reports since 2001 have documented problems in determining fair and reasonable prices within the Department of Defense (DoD) and federal contracting generally.

In July of 2014, another DoDIG report to the Director Defense Logistics Agency reviewed the prices for a sole-source commercial purchase. The DoDIG found that the contracting officer did not sufficiently determine whether prices were fair and reasonable for sole-sources spare parts negotiated for helicopters. The report found that the contracting officer used a previous purchase price without determining the reasonableness of that previous price and did not sufficiently review the contractor’s sales data in a noncompetitive environment. The report stated that the DLA potentially overpaid $9 million for this contract and may overpay as much as $2.6 million on future orders under this contract (DoDIG, 2014). The audit recommended that the DLA establish a quality assurance process that reviews whether contracting officers verify and document that sufficient analysis was performed to determine if the previous price was justified in accordance with DFARS PGI 215.403-3(4). The report was also directed to the Director of Defense Pricing. The report recommended that the Director of Defense Pricing issue guidance to establish a percentage of commercial sales that is sufficient to determine fair and reasonable prices when items are being acquired on a sole-source contract and market-based prices are used. The guidance should also require contracting officers to request information “other than certified cost or pricing data” to include cost data if sales data are not sufficient (DoDIG, 2014).

Gera and Maddox made similar recommendations in their October 2013 report. They suggested that the DoD implement oversight procedures to ensure the price analysis is documented and reviewed for completeness and adequacy and that local activities include internal controls to make sure price analysis is being done properly. In addition, they both forecasted that when price analysis is done poorly, it could cost us thousands, if not millions, in higher prices, such as the $9 million overpay found by the DoDIG in 2014 (Gera & Maddox, 2013, p. 51).

Previous research has been conducted on this subject outside of government audit reports. The findings from the research reports are consistent with the findings in government audit reports and do not appear to be surprising revelations.

As briefly discussed, there are multiple findings of deficiencies in each of the inspection reports; however, there were five deficiencies that repeated themselves more than others. First, the lack of proper price reasonableness determination documentation in the contracting file was mentioned in six separate inspection reports. Second, the failure to properly challenge commercial item designations, and third, the failure to properly verify previous prices paid as fair and reasonable, were both mentioned in four different reports. Fourth, the audit reports suggest that the wording of cost or pricing data regulations for commercial items has led contracting officers to hesitate to ask, or fail to ask, for certified cost and pricing data. Fifth, high workload and the amount of qualified contracting personnel were mentioned in four different reports.

In addition, the current authors highlighted three of the previously identified deficiencies (in Gera & Maddox, 2013). We then found that documentation of price reasonableness, and the failure to verify previous prices paid, were both apparent during the review of contract files during the research. The “personnel survey” results utilized during
the research identified that supervisors believed that manpower shortages were a reason for the failure to properly establish and document price reasonableness. Therefore, the recent research demonstrates that there is a consistency between the research findings and the inspection reports.

The existing literature provides evidence that this problem of conducting effective price reasonableness determinations will not be solved in the short term. Further research on these concerns is warranted.

Methodology

**Contract File Review**

We utilized a checklist for the contract file review. This checklist identified the price of each contract, commercial item designation, supply or service, estimates, competition, procedures, documentation present, and the price analysis techniques used. The parameters for file review were an acquisition price greater than $150,000 and under two years old. The DoD contracting office pulled a listing of contracts that met these parameters. Based on the contract list, we pulled randomly-selected files for review. The authors pulled a random sample of 30 contracts to provide the required data, and later we pulled another 20 contracts for a total of 50 contracts to be reviewed. The contract file sample contained 50 contract files valued at $72.2 million. They were a mix of supplies and services contracts. There were 19 contracts for the purchase of supplies valued at $23 million, and 31 contracts for services valued at $49.2 million. In addition, 41 of the 50 contracts (82%) in the sample were classified as commercial item contracts, and nine were classified as non-commercial; 14 contracts of the 41 (34%) were missing a statement in the file documenting that the item met the commercial item definition.

**Personnel Survey**

We utilized a personnel survey to anonymously determine the skill level and knowledge for conducting price analysis and price reasonableness determinations. Contracting personnel possess different skills and knowledge depending on years of experience, certification level, and the types of procurements they are typically assigned. The survey was designed for the participants to assess their knowledge base and provide feedback on how often they utilize various pricing techniques. The personnel survey was limited to the contracting personnel who worked in the same office that the contracting files were reviewed in order to compare results of the personnel survey to the contract file review. There were 25 surveys completed out of a possible 50 personnel who met the qualifications described in Section III: Price Reasonableness Determination. One survey respondent completed the demographic information only, and 24 survey respondents completed the entire survey. There were two supervisors that completed the supervisor section of the survey. The research was based on 25 survey respondents for the demographic information, and 24 respondents for the specific activity and knowledge-base questions. The 25 survey respondents who completed the demographic section of the survey indicated that 18 people (72%) had greater than five years of acquisition and contracting experience, six people (24%) had three to five years of experience, and one person (4%) had less than three years of experience. None responded that they had less than one year of experience.

**Comparisons Between the Contract File Data and Contract Survey Data**

This section makes comparisons of the file data and the personnel survey data that were directly related. Comparisons are necessary in order to relate the two data sets to each other in order to draw conclusions and see whether the perceptions of the operators match with the reality we see in the random files.
There was some discussion of competition in the pricing memorandums in 23 of the 50 contract files sampled. Pricing memorandums also indicated that one or more of the FAR 15 price analysis techniques were utilized on 48 of the 50 (96%) of the contract actions sampled. Of the 48 pricing memorandums, only 27 (56%), were justified by our examination. When the 24 survey respondents were asked, “Do you execute a pricing memorandum (or something similar) in the corresponding contract file which explains how you determined the offered and awarded price (or modified price) as fair and reasonable?”, 21 people (88%) answered that they frequently executed a pricing memorandum, 19 people (79%) responded that a pricing memorandum was critical, and 19 people (79%) responded that they had advanced or expert proficiency in completing pricing memorandums. The use of quantitative techniques (price indexing) was found only twice in the sample; however, a majority indicated usage of quantitative techniques when determining price reasonableness. Supervisors where split on the skill level of their employees in this matter.

There were 21 of 50 contract file actions that indicated that an IGCE was utilized for establishing price reasonableness; only five of those were substantiated by catalogs, contact with a vendor, or by a government technical report. When the 24 survey respondents were asked, “Does the customer provide an IGCE?”, seven people (29%) indicated that the requiring activity seldom/never provided an IGCE; seven people (29%) indicated that the requiring activity frequently provided an IGCE, and nine people (38%) indicated the customer always provided an IGCE.

When the 24 survey respondents were asked, “If the customer provides an IGCE, is it substantiated?”, 13 people (54%) indicated that it was seldom or never substantiated. When the 24 survey respondents were asked about their ability to determine whether an IGCE was reliable, 16 people (67%) of the respondents indicated yes. When the 24 survey respondents were asked whether IGCEs were used to determine price reasonableness for services, 11 people (46%) indicated yes.

A majority of contract files (25 of 32) in the sample contained market research information as required. The 18 contract actions were task/delivery orders where market research had occurred on the base contract and therefore were not counted. A majority of the contracting personnel surveyed indicated that market research was frequent (18 people), critical (18 people), and that contracting personnel were proficient at conducting market research (19 people). In a majority of the cases, contracting personnel indicated that both the customer and the contract specialist both provided market research information; however, 13 people (54%) questioned the reliability of customer-provided information. In 12 of the 25 research reports in the contract files pricing information in the market research was used for a price comparison.

Research Questions and Findings

We present answers to the research questions that drove this research by detailing applicable findings questions mentioned with our analysis of such. We follow these answers with our recommendations to the questions. In addition, the last two sections of this summary chapter include a discussion of the significance of the data we sampled for this report in comparison to what we found in our 2013 report, plus our suggestions for future research.

1. Do pricing memorandums deviate from Federal Acquisition Regulation (FAR), Defense Federal Acquisition Regulation Supplement (DFARS) requirements, and DFARS Procedures, Guidance and Information (PGI) procedures? There are multiple findings regarding this question. Rather than grouping our findings into one answer, we have addressed them individually (see Table 1).
Table 1. Summary of Inadequately Justified Price Analysis Documentation in the Files by FAR Price Analysis Technique

<table>
<thead>
<tr>
<th>FAR Price Analysis Techniques</th>
<th>Total</th>
<th>Supplies</th>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inadequate price competition</td>
<td>10 of 23</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Acceptance of prior prices without establishing their reasonableness</td>
<td>7 of 12</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Incomplete statements based on references to market research</td>
<td>0 of 12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Incomplete references to current price list, catalog, or advertisement</td>
<td>1 of 6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Incomplete comparison with IGCE or use of unreliable IGCEs</td>
<td>16 of 21</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Offeror did not provide data that was appropriate</td>
<td>0 of 1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No documentation in file for price reasonableness justification</td>
<td>2 of 50</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Totals of inadequate price analysis documentation for price reasonableness justification</td>
<td>36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1a. Inadequate Documentation Finding: A number of contract files that we reviewed were unable to demonstrate that prices paid were reasonable due to inadequate FAR price analysis methods, as depicted in Table 1: Summary of Inadequately Justified Price Analysis Documentation in the Files by FAR Price Analysis Technique.

1a. Analysis: From this data, we determined that the personnel involved in performing these contract actions did not include sufficient documentation to support the price analysis method used as required by FAR and DFARS. Considering the number of inadequate price analysis found in our sampled memos, it appears that contracting personnel are not familiar with how to appropriately perform and document price analysis. In particular, two types of price analysis, prior prices and IGCEs were performed and documented incorrectly more than 50% of the time. See details specific to prior price and IGCEs as follows under answers for 1c and 3b.

1a. Recommendation: A lack of supporting documentation could be easily corrected upon discovery by internal audits and returning the contract actions to have the applicable contracting personnel include the necessary documentation in the file. Additionally, reinforcement by internal procurement leadership in the importance of file documentation and including more hands on type training to make sure personnel understand what is and what is not proper supporting documentation.

We came across impressive recommendations from the contracting personnel who completed the surveys. When asked for suggestions on improving pricing documentation within the agency, respondents made the following comments:

- “It would be a good idea to review the FAR definitions of what constitutes an acceptable fair and reasonable determination. Even though training has been conducted on this topic, this is the sort of thing that should be reviewed periodically.”
• “A peer review could definitely be useful.”
• “(Perform) in-house quality assurance surveillance assessment (QASA) reviews (on the contract files) and provide training from the findings.”

When asked to make suggestions on DAU pricing courses, respondents replied with the following that support improved documentation:

• “Having a ‘real’ contract to analyze would definitely improve the courses.”
• “More on-site instruction. Not enough emphasis on sole-source/single source price analysis of commercial items. Most of the courses revolve around cost analysis—which is a small percentage of our buys. By the time we get an acquisition that requires cost analysis we’ve forgotten the course material (since most of the classes are taken within our 1st 2 years of employment).”
• “Often the applicability and level of documentation is missed.”

1b. Contract Review Board Checklist Finding: The contracting office for the DoD activity reviewed includes a local contract review board (CRB) checklist that also served as a business clearance memorandum (BCM) for significant actions as designated by the senior procurement officer. This checklist was extensive, including a section on pricing, and provided contracting personnel a means to check off any FAR/DFARS requirements, pre-award administration policies, procedures, and techniques that were utilized. The use of this checklist as a BCM makes it difficult for contracting personnel to fail to at least identify the type of price analysis utilized in an award decision.

1b. Analysis: We agree that the type of price analysis documented was indeed identified throughout the file review. It was the lack of poor documentation in the files that influenced our decisions to rate a pricing memo as justifiable or not. However, one would think that if the CRB were looking in depth into the documentation behind the identification of price analysis conducted, they would unearth some of the very same findings we made. Overall, we believe this checklist could be of use agency wide with the precaution of a deeper review of the documentation identified in the checklist.

1b. Recommendation: Well written checklist are indeed helpful to both contracting personnel and file reviewers. In particular, we found that file reviewers use the checklist without verifying the applicable documentation. Reviewers should build their own checklist based on the actual documentation they find and review instead of just using the existing checklist in the file for their review.

1c. Comparison of Current Offered Pricing Findings: Eleven pricing memorandums included some comparisons of current proposed (offered) prices when the lowest price was smaller than 80% of the next lowest price. For example, if the prices are $10, $50, and $55, respectively, then lower than 80% would be anything lower than 80% x $50, which equals $40. So the lower quote of $10 would be considered smaller.

In one file a price from a technically unacceptable offeror was still used to make a price comparison. This inappropriate comparison raises the issue that although competition is present and sought, is there actual price competition?

1c. Analysis: Proposed prices that are not within 80% of the next lowest price raise questions to the reliability of the proposed prices, and the existence of actual price competition. This could indicate that there is a mistake in offered price, a misunderstanding of the contract requirements, etc. In few cases there was documentation that included some determination of why there was such a large gap between the lowest price and next valid price or a price verification request by the CO to the lowest offeror.
According to the FAR Part 6, the award of a contract to a supplier based on lowest evaluated price alone can be false economy if there is subsequent default, late deliveries, or other unsatisfactory performance resulting in additional contractual or administrative costs. While it is important that government purchases be made at the lowest price, this does not require an award to a supplier solely because that supplier submits the lowest offer.

1c. Recommendation: A contractor who cannot perform is never a good deal at any price. More affirmative action needs to be taken by contracting personnel to affirm an offeror’s ability to perform at a low price which is considerably lower than the next offered price to ensure that the low price is fair and reasonable.

1d. Comparison of Proposed Price to Previous Price (Historical) Findings: Comparison of proposed prices to previous (historical) prices paid was utilized on 12 contract actions. In the seven instances that had an invalid previous price documented in the file, there was a previous price that could not be validated for one or more of the reasons displayed in Figure 1. Invalid previous (historical) prices were found because of one or more of the following reasons: time lapse, changes to terms and conditions, or uncertain reasonableness of the prior price.

1d. Analysis: If invalid previous prices are utilized, then price reasonableness has not actually been determined. For example, it would not be sufficient to use price(s) from a database produced by another contract specialist without understanding the type of analysis that was performed to determine the price. DoD-strengthened guidance on this subject is in PGI 215.403-3(4).

Since previous price comparison is one of the two preferred price analysis techniques, it is used quite often by contracting personnel in determining price reasonableness. This method is effective provided the validity of the comparison (similar items, categories, quantities, quality, qualifications, and/or circumstances) and the reasonableness of the previous price(s) can be established.

In this sample, more than 50% of the previous price comparisons made were invalid since the previous price was not verified. This illustrates why the authors determined that the contracts sampled do deviate from FAR/DFARS/PGI requirements and procedures. Further, if contracting personnel are not diligent in validating previous prices prior to using them for current pricing actions, then unreasonable prices can continuously perpetuate themselves into future contracting actions.

1d. Recommendation: Contracting personnel must first validate previous prices before using them in a price reasonableness determination and then adjust those prices to
make them comparable with the current offered price. If the previous price is not found to be valid, the contract specialist or contracting officer must find another method to determine price reasonableness. In-house reinforcement of how a previous price is verified is needed in this contracting activity. The DAU must include more classroom focus on this subject. The authors plan to write an article to be included in the National Contract Management Association Journal to assist in spreading the importance of performing this type of price analysis correctly.

Also, consider the following response to the DoDIG by the Director Defense Pricing for such situations. He suggested that contracting personnel use the support of the Pricing Centers of Excellence and the Contract Business Analysis Repository Information (CBAR) database. He stated that both methods are designed to supplement and improve the pricing skills of the acquisition workforce (DoDIG, 2014).

The authors agree that the Defense Director recommendation to use the CBAR database is useful, but its focus is on contract pricing actions exceeding $25 million. Also, the CBAR database does not include contractor sales data or validity of previous prices, which is what the DoDIG report number 2014-088 referred to.

2. Do pricing memorandums document the type of price analysis used in pricing formulation? What price analysis methods are being used?

**Findings:** All of the pricing memos documented some type of price analysis used in determining that the price was reasonable. The research findings show that 23 of the files used current competitive prices as a price analysis method (which is 46% of the total files). Comparison with the IGCE was documented in 21 pricing memos out of the 50 files, namely, 42% of the files. Previous prices (historical) documentation and comparison through market research were present in 12 of the 50 files, totaling 48% of the files as seen in Figure 2.

![Price Analysis Techniques Used](image)

**Figure 2. Contract File Data—Price Analysis Techniques Used**

**Analysis:** Current offered prices led the type of price analysis techniques used; IGCEs came closely behind, according to the contract files sampled. It is noteworthy that the contract files documented that adequate price competition through current offered prices is prevalent in most of the reviewed files. Normally, adequate price competition establishes price reasonableness and no other price analysis techniques are needed. IGCEs are not as
reliable as current offered prices so their usefulness is questioned. Further price analysis techniques should supplement an IGCE.

The application of price analysis techniques is notable. Contracting personnel at this office recognize the importance of price analysis in determining price reasonableness. However, as discussed in Question 1a (Inadequate Documentation), contracting personnel did not include sufficient documentation to support the price analysis method used as required by FAR and DFARS. Without the proper supporting documentation, the value of the techniques used are questionable.

Recommendation: Clearly the routine use of price analysis techniques is certainly present in pricing memos so recommendation is required. However, the recommended solutions for inadequate documentation under 1a apply at this point.

3. Do pricing memorandums refer to market research information or IGCE information? The findings regarding market research and IGCEs are not grouped together and we have addressed them individually, as follows:

3a. Market Research Findings: There were 12 contract pricing memorandums reviewed in the contract files that did utilize market research reports to establish price reasonableness, and a majority of the files in the sample contained market research reports; namely, of the 27 market research reports, 18 (67%) of those contracts addressed the type of pricing data collected, as shown in Figure 3. There were nine contracts in the sample that did not address the type of pricing data collected in the market research report, and there were five contracts in the sample that did not have a market research report that should have. In seven files, the IGCE and market research report were combined into one document.

![Figure 3. Contract File Data—Market Research Reports That Address Price](image)

When the 24 survey respondents were asked whether the customer provided market research information, 21 people (88%) responded that both the requiring activity and the contracting activity provided market research information. When asked about the adequacy of the market research information, 13 people (54%) indicated the customer seldom provided adequate market research information, nine people (38%) indicated that the customer frequently provided adequate market research information, and 8% of the respondents indicated the customer never provided adequate market research information.

3a. Analysis: The authors found that market research is included in most of the files we reviewed. Market research does improve the buyers’ understanding of pricing in the marketplace. The authors didn’t look in depth at the quality of the market research reports but did note that price was addressed in most of the market research reports examined. So
the authors conclude that market research reports, if reviewed, would have improved the buyers’ understanding of pricing in the marketplace. Redfern et al. (2013) found that contracting personnel did not always believe customer-provided information to be adequate; our research draws the same conclusion. According to the contracting personnel survey, 13 people (54%) indicated the customer seldom provided adequate market research information. Also, the authors found that reports that combined an IGCE with market research were confusing as to whether market research was done to support the IGCE or as full market research reports.

3a. Recommendation: Reliable market research from their customers/requirement activities will improve the buyers’ understanding of the marketplace. It will also give the buyer an opportunity to make better decisions when it comes to prices offered. The contracting and requiring activities should review agency guidance on market research and consider combined in house training. This will give both parties responsible for market research an opportunity to express their issues with conducting market research, documenting market research, and applying the information in the market research report to inform pricing determinations.

Gera and Maddox (2013) recommended that FAR Part 10 require that pricing be discussed in the market research reports (p. 54). The FAR identifies market research as a method for determining price reasonableness, but does not require that pricing be documented in the report. The authors still have the same opinion and would like to include a more extensive review of market research reports in the files.

For future research, the authors should review the quality of the market research reports. Knowledge of marketplace suppliers and prices can be critical to the government’s ability to negotiate a reasonable price. Poorly done market research could lessen an activity’s ability to achieve fair and reasonable prices. The authors would focus our research on how market research was conducted, and if it informed and influenced the contracting officer’s analysis of price reasonableness.

3b. IGCE Findings: The IGCE has two roles: one, as price analysis technique per FAR Parts 13 and 15, and two, IGCEs also support what the customer and contracting offices believes is the “should price” and should be completed before receipt of price proposals. In answering this question, the authors are only examining the IGCE’s role as a price analysis technique.

Next to current offered prices, IGCEs were heavily used as the basis for price reasonableness in 21 of the 50 files, essentially 42% of the contract actions reviewed. However, the authors found that only five of the 21 IGCEs used for determining price reasonableness could be substantiated and determined reliable. When the 24 survey respondents were asked, “If the customer provides an IGCE, is it substantiated?”, 13 people (54%) indicated that it was seldom or never substantiated. When the 24 survey respondents were asked about their ability to determine whether an IGCE was reliable, 16 people (67%) indicated yes. When the 24 survey respondents were asked whether IGCEs were used to determine price reasonableness for services, 11 people (46%) indicated yes.

Gera and Maddox (2013) recommended that agencies increase the importance of IGCEs. The contracting office we reviewed provided an IGCE template to the acquisition customer for assistance in filling out IGCEs since they are required to be completed by the acquisition customer. The IGCEs were required to be completed before receipt of the contractor’s price proposal. No further guidance on how complete the template was available. An IGCE was required on each purchase, but the template standard varied depending on whether the purchase was for supplies or services. A vendor quote would
satisfy the requirement for a commercial item supply purchase, but a more substantiated IGCE was required for services or non-commercial procurements.

The FAR lists an IGCE as a price analysis technique in both Parts 13 and 15, but it does not elaborate on what substantiates an IGCE. Best practices in pricing handbooks require that the source and methodology used in developing IGCEs be documented. More detailed information is found in CPRG Volume 1, paragraph 6.1.5, on techniques to properly validate and use an IGCE (Office of the Deputy Director of Defense, 2012).

3b. Analysis of IGCEs: The use of an IGCE to determine price reasonableness is frequent, and the reliability of IGCEs is not consistent. The contract personnel survey findings noted that more than 54% believed that IGCE are never substantiated by the customer but that 67% personally knew how to determine the reliability of an IGCE. Our actual contract file review findings raise questions regarding what the personnel in this survey considered to be a reliable IGCE. In contrast, only 23% of the IGCEs used in determining a fair and reasonable price could be substantiated as reliable. This real finding in the contract files contrasts with what contracting personnel perception that they know what is a reliable IGCE is. The fact that more than 54% questioned the reliability of a customer’s IGCE is discouraging because the customer’s IGCEs are still being used for determining price reasonableness.

We found that IGCEs were used as a primary price analysis comparison in 42% of the contract files we reviewed. In conclusion, this contracting office’s IGCEs basically are not effective in determining price reasonableness. However, the contracting personnel are still using the IGCE on a regular basis to determine a price is fair and reasonable though the contracting personnel know they are most likely not reliable. The fact that 77% of the IGCEs the authors examined are unreliable and, conversely, that 64% of the personnel think they know what a reliable IGCE is, indicates that the IGCEs in this contracting office’s files are truly questionable.

3b. Recommendations: The reliability of the IGCE should be considered when determining the usefulness of this technique. The authors recommend that contracting personnel become acquainted with what is a reliable IGCE by reviewing CPRG Volume 1, paragraph 6.1.5, on techniques to properly validate and use of an IGCE. DAU needs to add focus on the importance of the IGCE what makes it reliable in their pricing courses.

In addition, the authors still find last year’s recommendation in their 2013 report also applicable. The authors propose that government activities increase the importance of IGCEs and consider the following steps to make IGCEs more reliable for use in conducting price analysis: (1) the contract specialists should be presented with good training on what good IGCEs are and what to document, (2) the individuals that develop IGCEs need to know how to do it (there should be more specific training in this area), (3) an online check system should be put in place where government IGCEs are accepted if and only if the substantiation is provided, and (4) there should be consideration to acknowledge IGCEs in the FAR/DFAR/PGI with more importance than it currently is given. The policy-makers need to appreciate how much IGCEs are being used and how more guidance will assist s in determining what the contract specialist “should pay” and hopefully giving the contract specialists the ammunition to negotiate better prices (Gera & Maddox, 2013, p. 62).

4. If deviations in pricing memorandums exist, do they differ by the same characteristics and/or by unsimilar characteristics?
In 46% of the files, deviations in pricing memorandums do exist as depicted as inadequate justification in Table 2. Some deviations found in the files reviewed were consistently the same and others uniquely not similar to others.

Table 2. Contract File Data—Adequate vs. Inadequate Justification for Price Reasonableness

<table>
<thead>
<tr>
<th></th>
<th>Adequate Justification</th>
<th>Inadequate Justification</th>
<th>Documentation not in file</th>
<th>Total</th>
<th>Percent Justified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contract Files</td>
<td>27</td>
<td>21</td>
<td>2</td>
<td>50</td>
<td>54%</td>
</tr>
<tr>
<td>Services</td>
<td>15</td>
<td>16</td>
<td>0</td>
<td>31</td>
<td>48%</td>
</tr>
<tr>
<td>Supplies</td>
<td>12</td>
<td>5</td>
<td>2</td>
<td>19</td>
<td>63%</td>
</tr>
</tbody>
</table>

A majority of the pricing memorandums do deviate by two consistent characteristics: the lack of supporting documentation to justify the technique utilized to establish price reasonableness, and the use of unsupported IGCEs. See findings that support inadequate documentation under the answers to 1a and findings behind the unsupported IGCEs in the answers to Question 3b.

The pricing memorandums in the sample reviewed established that the contract specialist determined price reasonableness as well as listed the technique utilized; however, substantiating documentation (e.g., calculation sheets, reference materials such as catalog data found on line, copy of previous price documentation, and methodology) are not always included, supporting the source of their recommendations. A very small percentage of the IGCEs in the contract files were reliable enough to support that the offered price being analyzed was indeed reasonable. Contracting personnel listed the IGCE as justifying the price of 17 service and four supporting contracts. Only five of the 21 were judged as reliable and therefore substantiated by the authors.

Some of the files reviewed contained unique deviations in their pricing memos (Maddox, Fox, & Gera, 2014).

5. What are the most predominant price analysis techniques exercised in purchasing supplies versus services?

This data is listed in Table 3, and offers insight into the predominant type of price analysis techniques exercised in purchasing supplies versus services.
Table 3. Contract File Data—Answers to the Question “What Was the Documented Justification for Price Reasonableness?”

<table>
<thead>
<tr>
<th>Price Justification in Pricing Memos</th>
<th>Total number</th>
<th>Supply</th>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison to current offered price?</td>
<td>23</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Comparison to previous prices paid?</td>
<td>12</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Parametric estimating?</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Competitive published price lists?</td>
<td>6</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Comparison of proposed prices with independent government cost estimates?</td>
<td>21</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>Comparison of proposed prices obtained through market research?</td>
<td>12</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Analysis of sales data provided by the offeror?</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Findings:** According to the contract file data sampled, contracts for both supplies and service made use of comparisons to current offered prices and market research prices equally as price analysis techniques.

Services are far more dependent on IGCEs than supplies. We found that 17 of 31, or 55%, of service contracts tapped IGCEs as a primary price analysis technique. Only four of 19 supplies, 21%, benefitted from IGCEs.

Supplies take advantage of competitive price lists in justifying price reasonableness much more often than services. Five out of six employments of competitive price lists as a price analysis technique, represents a rate of 83% utilization for supplies versus services.

**Analysis:** The authors don’t know the why part of the answer, but can infer that it is much easier to find prices in the marketplace for supplies than for services, thus the use of competitive price lists. For services, there is more dependence on IGCEs to make price comparisons, since IGCEs generally disclose an estimate of labor hours by the type of effort required. IGCEs are more effective for justifying the price of services than other price analysis techniques outside of two or more current offered prices.

**Recommendations:** This is the same recommendation as last year’s report by Gera and Maddox (2013):

Buying services is different than buying supplies; that also means that they are different when it comes to pricing. A step should be added to the services acquisition guidebook that focuses just on the pricing of services. Possibly the FAR, DFARS, and PGI need to reframe price analysis methods that are more useful in purchasing services, as opposed to current references to supplies only.

One of the authors discussed the uniqueness of pricing services with Frank Kendall, U.S. defense undersecretary for acquisition, technology and logistics at 2014’s AFCEA symposium. He agreed and encourage the author to explore that concern. Unfortunately she is no longer instructing in contract pricing. She is now holding employment that little allows
this type of activity nor the access to students and peers in generating new ideas. A review of how the commercial world prices services would have been the place to start.

IGCEs play such an important part in pricing services as a “should price” before the request for vendor prices and as a tool to determine price reasonableness after a price is proposed. The generation of local guidance and policies and actual instruction to educate the customer/requirements personnel could bring true and valuable savings.

6. Why do pricing memorandums lack sufficient justifications and supporting information? What challenges are present in executing price analysis?

Findings: When the 24 survey respondents were asked what the challenges in executing price analysis, determining price reasonableness, and documenting were, four people (17%) indicated a lack of knowledge in conducting price reasonableness determinations, nine people (38%) indicated the amount of time to complete price reasonableness determinations, four people (17%) gave other reasons as challenges, and seven people (29%) did not answer the question (see Figure 4).

Figure 4. Answers to the Question “What Is Your Challenge in Executing Price Analysis, Determining Price Reasonableness, and Documenting?”

The respondents who indicated other reasons gave the following inputs:

- “Lack of competition and sufficient IGCEs can be somewhat challenging.”
- “Sole-source acquisitions—finding similar products in the market to compare prices with.”
- “Conducting market research when the item is sole source/single source and no other vendors can provide price quote.”
- “Inadequate product description from the requiring activity.”
- “Lack of similar offerings; sole source prevents apples-to-apples price analysis with other offerings.”

Analysis: Redfern et al. (2013) found that pricing memorandums lack justifications and supporting information because of a lack of time to complete adequate price reasonableness determinations. Our research draws the same conclusion. A majority of the survey respondents who answered this question also gave specific reasons as to why time was a challenge. Each comment can be linked to specific time-related measures, such as procurement action lead times (PALT) or other time-consuming activities directly linked to the procurement.
The personnel survey results indicate that the contracting personnel in the surveyed office were well qualified, with a majority of personnel having more than five years of experience. Eighteen of 24 surveyed had both Level I and II pricing course, namely 75%. However, four personnel had no Level I pricing training. Price Analysis is only taught in Level I courses.

The survey did not ask respondents to quantify manpower shortages, but at least one supervisor did believe manpower shortages were a problem for the agency. The supervisor’s opinion that manpower shortages are a problem is consistent with the findings of several DoDIG reports that found high workloads and shortages of qualified personnel lead to the government paying more than what is fair and reasonable.

Another supervisor identified the skill level of the employees as the organization’s biggest challenge. Considering the poor quality of price analysis (unjustified) documentation found in 23 of 50, 46% of files reviewed, the authors agree and conclude that employees skill levels are not up to standard.

Overall, the use of price analysis techniques are common, but there are serious deficiencies when it comes to actually using the techniques correctly and including proper supporting documentation. Poor documentation to support the price reasonableness determination was the biggest weakness. When competition by itself does not establish price reasonableness, the most utilized techniques for determining price reasonableness within this office were comparison through market research, comparison to previous pricing, and comparison to IGCEs. The use of indexing, regression, and parametric analysis is uncommon for commercial item purchases; however, contracting professionals in this contracting office are aware of the techniques and are trained to use them. Consistent with DoDIG report findings, this contracting activity is concerned with high workloads and shortages of personnel. Survey results and comments show that contracting personnel are spending time either validating customer requirements or researching requirements on their own to validate inadequate customer provided market research information and customer-provided IGCE documentation.

The contracting office that provided the contract file data and personnel survey data is concerned with pricing inaccuracies and seeks to constantly improve its adherence to policy with extensive internal reviews. However, the number of unjustified pricing memos we reviewed is worrisome. The authors found that over $3 million in supplies and $39 million in services was not adequately justified for price reasonableness.

**Recommendation:** This activity needs to take a hard look at what is disabling the personnel in their activity from performing price analysis properly, such as

- Determine if their current assessment methods like the contract review board and quality assurance assessment team consistently follow price reasonableness standards in accordance with the FAR/DFARS.
- Train and retrain contracting personnel on price analysis techniques in determining price reasonableness along with what is proper support documentation for pricing.
- Eliminate or reduce the challenges that contracting personnel have in executing proper price reasonableness. For example, guidance is needed on “conducting market research when the item is sole source/single source and no other vendors can provide price quote.”
- Add adequate guidance on the preparation of IGCEs and market research reports by customers (requiring activities).
Comparison to Past Research

Contract File Data

It would be easy to just go table by table and list the differences and the similarities, but we want to focus on the real reason why we are doing this research and what we have learned so far. Basically, the authors wanted to know if price analysis and the associated price analysis techniques are being exercised and conducted correctly in determining price reasonableness. The authors believe that when price analysis is not conducted properly, there is reason to believe the prices offered and paid are unreasonable.

Our data in this report came from a different agency. Dollar values were higher (nothing below $150,000), and there appeared to be much more oversight of buys at any value within this agency. Little use of Part 13 simplified pricing procedures were found, like last year’s data in purchasing commercial items. However, with all of these differences, the results are very much the same, such as

- Inadequate documentation
- Little use of quantitative skills learned in pricing classes
- Unreliable IGCEs
- Infrequent requests for offeror data when needed

This year, the authors added emphasis on auditing previous prices as a price justification and particularly if they were sole source. In many of the contract files reviewed, the authors found that contract specialists were using previous prices that were not reasonable and continuously perpetuating the same unreasonable price for justification on other proposed price. The authors also prodded further into the documentation that used current offered prices as the price justification. The authors found that contracting personnel were not verifying low prices that were considerably less than the next offered price. It would have been interesting to find out if the contractor in those cases actually delivered at such low prices. In reviewing what type of price justifications are used in supplies versus services, the authors found few differences. Essentially, in cases other than current offered prices, IGCEs were the most used for justifying a price for services and published price lists/market research for supplies. One significant improvement in our current research was the access to real paper files. In 2014, the current authors had only access to electronic files. It was hard to retrieve and review files if they were not well filed. Having paper files to review allowed the authors to now look more closely at the documentation, and it was much easier to find documentation misfiled or misnamed.

The authors’ recommendations are not very different. They are similar issues, same suggestions. The authors would like to conduct a comprehensive exploration of market research reports. The authors think there may be a link to why so many IGCEs are unreliable.

Inaccuracies in pricing memorandums can result in increased procurement costs, sustained protests, and loss of agency contracting authority, and the authors conclude that those inaccuracies continue to occur at most, if not all, DoD agencies. The big questions is, can we remove the reasons for these inaccuracies?

Personnel Survey Data

In comparing the current research surveys to last year’s report, one would think this is the same personnel who submitted their responses. The majority of the personnel have more than five years of experience in supplies or services. There are few responses from personnel with three or fewer years’ experience. The majority were qualified at Level II. Very
few were at Level III. Most had taken Level I and II contract pricing courses. “Time to
complete” and “inadequate skill level” were still the challenge in executing price analysis
according to the nonsupervisory personnel. In both years, supervisors pointed to skill level
as an issue. In comparing the file data and survey data, personnel responded that they use
quantitative methods often, but the authors did not find any real use of quantitative methods
in the contract file pricing documentation.

The authors’ added new questions to the survey that directly asked the survey
participant to rate the effectiveness of DAU contract pricing classes. The majority said they
were effective. However, all the written comments by the surveyed participants supported
that price analysis is underemphasized in the DAU courses. In potential future research
efforts, the researchers will consider scaling down the personnel survey questions and focus
on the areas of price analysis where the authors historically found inadequate
documentation, and explore survey participants’ opinions on why the use of IGCEs are more
prevalent in service versus supply purchases.

References
Office of Federal Procurement Policy and the United States Congress. Retrieved from
http://www.acquisition.gov/comp/aap/24102_GSA.pdf
Defense Federal Acquisition Regulations Supplement (DFARS)/Procedures, Guidance, and
3243 (1994).
http://www.federaltimes.com/article/20140826/BLG06/308260017
Gera, R., & Maddox, J. (2013). Price analysis on commercial items purchases within the
Department of Defense. In Proceedings of the Eleventh Annual Acquisition Research
GAO. (2006). Contract management: DoD vulnerabilities to contracting fraud, waste, and
Maddox, J., Fox, P., & Gera, R. (2014). Price analysis on commercial item purchases within
the Department of the Navy (Sponsored report). Monterey, CA: Naval Postgraduate
School, Acquisition Research Program.
Office of the Deputy Director of Defense Procurement and Acquisition Policy for Cost,
within the Department of Defense (Joint Applied Project, Naval Postgraduate School).
Retrieved from http://calhoun.nps.edu/public/handle/10945/37743
Management Concepts.
Panel 20. Understanding the Relevance of Big Data and Predictive Analytics in Business and the Defense Industrial Base

Thursday, May 14, 2015

3:30 p.m. – 5:00 p.m.

Chair: Mark Krzysko, Deputy Director, Acquisition Resources and Analysis, Office of the Under Secretary of Defense (Acquisition, Technology, & Logistics)

Lexical Link Analysis Application: Improving Web Service to Acquisition Visibility Portal Phase III

Ying Zhao, NPS
Doug MacKinnon, NPS
Shelley Gallup, NPS

Mark Krzysko—serves as the deputy director of the Enterprise Information and Office of the Secretary of Defense Studies. In this senior leadership position, he oversees Federally Funded Research and Development Centers and directs data governance, technical transformation, and shared services efforts to make timely, authoritative acquisition information available to support oversight of the Department of Defense’s major programs—a portfolio totaling more than $1.6 trillion of investment funds over the life cycle of the programs.

Preceding his current position, Krzysko served as ADUSD for business transformation, providing strategic guidance for re-engineering the Department’s business system investment decision-making processes. He also served as ADUSD for strategic sourcing & acquisition processes and as director of the Supply Chain Systems Transformation Directorate, championing and facilitating innovative uses of information technologies to improve and streamline the supply chain process for the Department of Defense. As the focal point for supply chain systems, he was responsible for transformation, implementation, and oversight of enterprise capabilities for the acquisition, logistics, and procurement communities. In addition, Krzysko served as advisor to the deputy under secretary of defense for business transformation on supply chain matters and as the functional process proponent to the Department’s business transformation efforts, resulting in the establishment of the Business Transformation Agency.

In March 2002, Krzysko joined the Defense Procurement and Acquisition Policy office as deputy director of e-business. As the focal point for the acquisition domain, he was responsible for oversight and transformation of the acquisition community into a strategic business enterprise. This included driving the adoption of e-business practices across the Department, leading the move to modernize processes and systems, and managing the investment review process and portfolio of business systems. Krzysko served as the division director of Electronic Commerce Solutions for the Naval Air Systems Command from June 2000 to March 2002. From April 1991 until March 2000, Krzysko served in various senior-level acquisition positions at the Naval Air Systems Command, including contracting officer of F/A-18 foreign military sales, F/A-18 developmental programs, and the F-14. In addition, he served as program manager of Partnering, the Acquisition Business Process Re-engineering Effort, and as acquisition program manager for the Program Executive Office for Tactical Aircraft.

Krzysko began his career in the private sector in various executive and managerial positions, including assistant managing director for Lord & Taylor Department Stores and operations administrator for Woodward & Lothrop Department Stores. Krzysko holds a Bachelor of Science
degree in finance from the University of Maryland University College, College Park, MD, and a Master of General Administration degree in financial management from the same institution.
Lexical Link Analysis Application: Improving Web Service to Acquisition Visibility Portal Phase III

Ying Zhao—is a research associate professor at the Naval Postgraduate School and frequent contributor to DoD forums on knowledge management and data sciences. Her research and numerous professional papers are focused on knowledge management approaches such as data/text mining, Lexical Link Analysis, system self-awareness, Collaborative Learning Agents, search and visualization for decision-making, and collaboration. Dr. Zhao was principal investigator (PI) for six contracts awarded by the DoD Small Business Innovation Research (SBIR) Program. Dr. Zhao is a co-author of four U.S. patents in knowledge pattern search from networked agents, data fusion, and visualization for multiple anomaly detection systems. She received her PhD in mathematics from MIT and is the co-founder of Quantum Intelligence, Inc. [yzhao@nps.edu]

Doug MacKinnon—is a research associate professor at the Naval Postgraduate School. MacKinnon is the deputy director of the Distributed Information and Systems Experimentation (DISE) research group where he leads multidisciplinary studies ranging from leading the Analyst Capability Working Group for the U.S. Air Force and studying Maritime Domain Awareness, as well as Knowledge Management (KM) and Lexical Link Analysis projects. He also led the assessment for the Tasking, Planning, Exploitation, and Dissemination process during the Empire Challenge 2008 and 2009 (EC08/09) field experiments and for numerous other field experiments of new technologies during Trident Warrior 2012. Dr. MacKinnon teaches courses in operations research and holds a PhD from Stanford University, conducting successful theoretic and field research in KM. He has served as the program manager for two major government projects of over $50 million each, implementing new technologies while reducing manpower requirements. He has served over 20 years as a naval surface warfare officer, amassing over eight years at sea and serving in four U.S. Navy warships with five major underway deployments. [djmackin@nps.edu]

Shelley Gallup—is a research associate professor at the Naval Postgraduate School’s Department of Information Sciences and the director of Distributed Information and Systems Experimentation (DISE). Dr. Gallup has a multidisciplinary science, engineering, and analysis background, including microbiology, biochemistry, space systems, international relations, strategy and policy, and systems analysis. He returned to academia after retiring from naval service in 1994 and received his PhD in engineering management from Old Dominion University in 1998. Dr. Gallup joined NPS in 1999, bringing his background in systems analysis, naval operations, military systems, and experimental methods first to the Fleet Battle Experiment series (1999–2002) and then to Fleet experimentation in the Trident Warrior series (2003–2013). Dr. Gallup’s interests are in knowledge management and complex systems field experimentation. [spgallup@nps.edu]

Abstract

We have been studying DoD acquisition decision-making since 2009. The U.S. DoD acquisition process is extremely complex. There are three key processes that must work in concert to deliver capabilities: determining warfighters’ requirements/needs, the DoD budget planning, and the procurement of final products. Each process produces large amounts of information (Big Data). There is a critical need for automation, validation, and discovery to help acquisition professionals, decision-makers, and researchers understand the important content within large data sets and optimize DoD resources throughout the processes. Lexical Link Analysis (LLA) can help, by applying automation to reveal and depict—to decision-makers—the correlations, associations, and program gaps across all or subsets of acquisition programs over many years. This enables strategic understanding of data gaps and potential trends, and can inform managers where areas might have higher program risk and how resource and big data management might affect the desired return on investment among projects. In this paper, we describe new developments in analytics and visualization, how LLA is adaptive to Big Data Architecture and Analytics (BDAA), and needs for Big Acquisition Data used in Defense Acquisition Visibility Environment (DAVE).
Background

We have been studying Department of Defense (DoD) acquisition decision-making since 2009 (Gallup et al., 2009; Zhao, Gallup, & MacKinnon, 2010, 2011, 2012a, 2012b, 2013, 2014). The U.S. DoD acquisition process is extremely complex. There are three key processes that must work in concert to deliver capabilities: definition of warfighters’ requirements/needs, DoD budget planning, and procurement of products, as in Figure 1. Each process produces volumes of information (Big Data). The need for automation, validation, and discovery is now a critical need, as acquisition professionals, decision-makers, and researchers grapple to understand data and make decisions to optimize DoD resources.

**Figure 1. DoD Acquisition Decision-Making**

Since 2009, we have been working on the problem of how the interlocking systems processes become aware of their fit between DoD programs and warfighters’ needs. How are gaps revealed? Moreover, in the performance of DoD acquisition processes, each functional community is required to review only the particular information for which it is responsible, further exacerbating the problem of lack of fitness. For example, the systems engineering community typically only examines the engineering documents and feasibility studies, the test and evaluation community looks only at the test and evaluation plans, and the acquisition community looks at the acquisition strategies. Rarely do these stakeholders review each other data or jointly discuss the core questions and integrated processes together as shown in Figure 1.

Motivated by this lack of fit and horizontal integration, we have been applying Lexical Link Analysis (LLA), a data-driven automation technology and methodology across DoD acquisition processes to

- surface themes and their relationships across multiple data sources
- discover high value areas for investment
- compare/correlate data from multiple data sources
- sort/rank important and interesting information

LLA is a data-driven method for pattern recognition, anomaly detection, and data fusion. It shares indexes not data, feasible for parallel and distributed processing, adaptive to Big Data Architecture and Analytics (BDAA) and needs for Big Acquisition Data.
As an example from past work, we took a detailed look at the Research, Development, Test and Evaluation (RDT&E) budget modification practice from one year to the next over the course of 10 years and about 450 DoD program elements. We found a pattern that the programs with fewer links (measured by LLA) to warfighters’ requirements received more budget reduction in total but less on average, indicating the budget reduction may have focused only on large and expensive programs rather than perhaps cutting all the programs that do not match warfighters’ requirements. Furthermore, the programs with more links to each other received more budget reduction in total, as well as on average, indicating a pattern of good practice of allocating DoD acquisition resources to avoid overlapping efforts and to fund new and unique projects. These findings were useful as validation and guidance for future decision processes for automatically identifying programs to match the warfighter’s requirements, limit overall spending, minimize efficiencies, eliminate unnecessary cost, and maximize the return on investment.

In this paper, we demonstrate a set of comprehensive LLA analysis reports and visualizations generated automatically from multiple data sources. These reports and visualizations reveal data correlations and gaps among multiple data sources. These correlations and gaps could form the basis for pattern recognition, anomaly detection, and further inquiry or future reconciliation of the expectations (e.g., acquisition strategy) and realities (e.g., engineering feasibility) from various communities. The automatic discovery of the disconnection or gaps could be fed back to the human analysts or decision-makers for decision-making and resource management.

Methodology

**Lexical Link Analysis (LLA)**

LLA has been used to analyze unstructured and structured data for pattern recognition, anomaly detection, and data fusion. It uses the theory of system self-awareness (SSA) to identify high-value information in the data that can be used to guide future decision processes in a data-driven or unsupervised learning fashion. It is implemented via a smart infrastructure named “system and method for knowledge pattern search from networked agents (U.S. patent 8,903,756),” also known as Collaborative Learning Agents (CLA), licensed from Quantum Intelligence, Inc. (Zhou, Zhao, & Kotak, 2009).

In LLA, a complex system is expressed in specific vocabularies or lexicons to characterize its features, attributes, or surrounding environment. LLA uses bi-gram word pairs as the features to form word networks. Figure 2 depicts using LLA to analyze 10 years of reports in the Naval Postgraduate School (NPS) Acquisition Research Program with word pairs as groups or themes. Figure 3 shows a detail of a theme in Figure 2. A node represents a word. A link or edge represents a word pair.
LLA is related to bags-of-words (BAG) methods such as LDA (Blei, Ng, & Jordan, 2003) and text-as-network (TAN) methods such as the Stanford Lexical Parser (SLP; Stanford Natural Language Processing Group [SNLPG], 2015). LLA selects and groups features into three basic types:

- **Popular (P):** They are the main themes in the data. Figure 3 is an example of a popular theme centered around word nodes “analysis,” “model,” and “approach.” These themes could be less interesting because they are already
in the public consensus and awareness. They represent the patterns in the data.

- Emerging (E): Themes may grow to be popular over time. Figure 4 is an example of an emerging theme centered around word nodes “national,” “defense,” and “acquisition.”

- Anomalous (A): These themes may be off-topic themes that are interesting for further investigation. Figure 5 is an example of anomalous theme centered around word nodes “stock” and “market(s).”

![Figure 4. An Example of Emerging Theme](image1)

![Figure 5. An Example of Anomalous Theme](image2)

Figure 6 summarizes LLA used for historical and new data. The red part shows a pattern (e.g., a theme) discovery phase using historical data including data fusion that come...
from multiple learning agents. The black part shows an application phase that new data is compared with the patterns discovered and hence the anomalies are revealed.

Figure 6. Diagram for the LLA Method

**Word Pairs Generalization and CCC Method**

Figure 7 shows the word pairs/bi-gram in an LLA can be generalized as a Context-Concept-Cluster (CCC) model, where a context is a generic attribute that can be shared by multiple data sources, a concept is a specific attribute for a data source, and a cluster is a combination of attributes or themes that can be computed using a word community finding algorithm (e.g., Girvan & Newman, 2002) in Figure 6 to characterize a data set. Context can be a word, location, time, or object, and so on.
Figure 7. The Word Pairs/Bi-Gram in LLA a Context-Concept-Cluster (CCC) Model

Figure 8 summarizes how a generalized CCC method is used for historical and new data. Similar to Figure 6, there is a pattern discovery phase using historical data where patterns are learned and discovered, and an application phase for a new data is compared with the patterns discovered, and anomalies are revealed.
Research Results

Task 1

We are working with the OSD OUSD(ATL) (US) to install the LLA/SSA/CLA system as a web service using a Linux platform (i.e., CentOS) in the Defense Acquisition Visibility Environment (DAVE) test bed. We created a publically available data set with the installation to test. In this example, data sources include 10 days of business news of about 1,000 companies, which are organized in industries as follows:

- Technology
- Services
- Healthcare
- Utilities
- Basic Materials
- Financial
- Consumer Goods
- Industrial Goods
- Conglomerates

Each category of information such as “Healthcare” or “Consumer Goods” are indexed, mined, and listed under “Index Management” separately in a single LLA server. When clicking “Fuse,” these indexed/mined models are fused into one model. Figure 9 shows “Fuse Results” from LLA listed.

![Index Management](image)

Figure 9. Fuse Results Listed

Figure 10 shows the discovered themes, where green themes 101(P) and 20(P) are “popular” themes, blue themes 156(E), 49(E), and 46(E) are “emerging” themes, and gold themes 208(A), 62(A), and others are “anomalous” themes.

- Popular themes are the main themes in the data. Figure 11 is an example of a popular theme centered “dividend cuts, see dividend” for this data. Columns “Consensus” is the ratio of the number of matched word pairs (i.e., at least two data sources contain the word pairs) over the number of unique word pairs (i.e., only one data source contains the word pairs). These themes could be less interesting because they are already in the public consensus...
and awareness. They represent the patterns in the data. The red links represent the word pairs that are shared for at least two data sources while the black data sources are unique to one data source.

- Emerging themes may grow to be popular over time. Figure 12 is an example of an emerging theme centered “back shares, Canada back.”
- Anomalous themes may be off-topic themes that are interesting for further investigation. Figure 13 is an example of anomalous theme centered around “top buys, set top.” Anomalous concepts are more interesting to investigate, for example, concepts in Figure 13 such as “buys web,” “streaming service,” “buys insider,” “web ipo,” and so on, may have better returns on investment than the concepts in a popular theme such as “sees dividend” and “announces positive.”

**Discovered Themes**

<table>
<thead>
<tr>
<th>Event Date Sort</th>
<th>Theme Id</th>
<th>Theme Keywords</th>
<th>Visualization</th>
<th>Matched</th>
<th>Unique</th>
<th>Total</th>
<th>Consensus/Gaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/10</td>
<td>DIVIDEND CUTS, SEES DIVIDEND</td>
<td>YES</td>
<td>164</td>
<td>27</td>
<td>191</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>10/10</td>
<td>ANNOUNCES POSITIVE</td>
<td>YES</td>
<td>21</td>
<td>2</td>
<td>23</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>BACK SHARES, CANADA BACK</td>
<td>NO</td>
<td>142</td>
<td>24</td>
<td>167</td>
<td>0.86</td>
<td>0.14</td>
</tr>
<tr>
<td>all</td>
<td>COMPANIES START/HELP/PS</td>
<td>NO</td>
<td>11</td>
<td>4</td>
<td>15</td>
<td>0.72</td>
<td>0.38</td>
</tr>
<tr>
<td>all</td>
<td>COMPANY-ANNOUNCED EXECUTIVE NAMED</td>
<td>NO</td>
<td>11</td>
<td>22</td>
<td>33</td>
<td>0.84</td>
<td>0.16</td>
</tr>
<tr>
<td>all</td>
<td>TOP BUYS, SET TOP</td>
<td>NO</td>
<td>128</td>
<td>32</td>
<td>160</td>
<td>0.80</td>
<td>0.20</td>
</tr>
<tr>
<td>all</td>
<td>EXPANDS, HEALTH, DATA, EXPAND</td>
<td>YES</td>
<td>80</td>
<td>14</td>
<td>94</td>
<td>0.81</td>
<td>0.19</td>
</tr>
<tr>
<td>all</td>
<td>PLANS CUSTOMER-IP PLANS</td>
<td>YES</td>
<td>17</td>
<td>17</td>
<td>34</td>
<td>0.82</td>
<td>0.18</td>
</tr>
<tr>
<td>all</td>
<td>SECURITIES, CALIFORNIA, MORTGAGE, SECURITIES</td>
<td>YES</td>
<td>76</td>
<td>18</td>
<td>94</td>
<td>0.81</td>
<td>0.19</td>
</tr>
<tr>
<td>all</td>
<td>CAPITAL FUND, FUND RATING</td>
<td>YES</td>
<td>101</td>
<td>20</td>
<td>123</td>
<td>0.84</td>
<td>0.16</td>
</tr>
<tr>
<td>all</td>
<td>CITY CENTER, EUROPEAN CENTER</td>
<td>YES</td>
<td>57</td>
<td>17</td>
<td>74</td>
<td>0.77</td>
<td>0.23</td>
</tr>
<tr>
<td>all</td>
<td>UPDATE SUNOCO, UPDATE IR</td>
<td>YES</td>
<td>66</td>
<td>9</td>
<td>55</td>
<td>0.84</td>
<td>0.16</td>
</tr>
<tr>
<td>all</td>
<td>HIGHLIGHTS CONOCOPHILLIPS, HIGHLIGHTS HSBC</td>
<td>YES</td>
<td>68</td>
<td>12</td>
<td>80</td>
<td>0.85</td>
<td>0.15</td>
</tr>
<tr>
<td>all</td>
<td>FOOD INDUSTRY, INDUSTRY LEADING</td>
<td>YES</td>
<td>38</td>
<td>12</td>
<td>50</td>
<td>0.76</td>
<td>0.24</td>
</tr>
<tr>
<td>all</td>
<td>LEAD, COUNTRY, SYSTEM LEAD</td>
<td>YES</td>
<td>3</td>
<td>7</td>
<td>10</td>
<td>0.82</td>
<td>0.18</td>
</tr>
<tr>
<td>all</td>
<td>WEEKLY MARKET, WEEKLY RECAP, MARKET RECAP</td>
<td>YES</td>
<td>41</td>
<td>8</td>
<td>49</td>
<td>0.88</td>
<td>0.12</td>
</tr>
<tr>
<td>all</td>
<td>CARD DEALS, GET, CARD, GET, CARD</td>
<td>YES</td>
<td>59</td>
<td>12</td>
<td>71</td>
<td>0.76</td>
<td>0.24</td>
</tr>
</tbody>
</table>
Figure 11. Visualization for the Popular Theme 101(P)

Figure 12. Emerging Themes (e.g., 156[E])
Figure 13. Anomalous Theme (e.g., 208[A])

**Match Matrix Unique Word Pairs by Theme**

Figure 14 shows the numbers of unique word pairs in a data source and a theme. For example, there are 12 unique word pairs for the data source “Index_BasicMaterials” in the theme titled “101:dividend cuts, sees dividend” in Figure 14. Clicking this number leads to a list showing the 12 word pairs as shown in Figure 15. Figure 16 shows the list can be further drilled down to a search result list (e.g., “sees energy”) and to the original documents that contain the word pair.

![Match Matrix Unique Word Pairs by Theme](image)

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Index_BasicMaterials</th>
<th>Index_Financial</th>
<th>Index_Commodities</th>
<th>Index_Commercial</th>
<th>Index_Power</th>
<th>Index_Auto</th>
<th>Index_Recreation</th>
<th>Index_Energy</th>
<th>Index_Consumer</th>
<th>Dividend Cuts</th>
<th>Dividend Growth</th>
<th>Dividend Rate</th>
<th>Dividend Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>80.00</td>
<td>20.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
<td>60.00</td>
</tr>
</tbody>
</table>

Figure 14. Match Matrix Unique Word Pairs by Theme
Figure 15. Drill-Down List for the Unique Word Pairs in Theme 101 and Index_BasicMaterial

Figure 16. Drill-Down Search on “Sees Energy”
**Match Matrix**

Figure 17 shows the match matrix for comparing data sources. The column “Match Score” shows the number of matched word pairs for Index_BasicMaterials. "5.00(0.02)" shows the number (5) of matched word pairs and correlation (0.02) between Index_BasicMaterials and Index_Financial. The correlation, computed as 

\[
\frac{5}{\sqrt{30+223}} \times \sqrt{28+270},
\]

is normalized using the match scores and uniqueness scores for both data sources. Clicking on the “5.00(0.02)” leads to the list of the matched word pairs for the two sources as shown in Figure 18. Clicking on “Energy Prices” or “Oil Fund” (i.e., the red boxes in Figure 18) leads to the search results of two terms respectively. The search results are sorted in a descending order of the counts of how many “popular,” “emerging,” and “anomalous” word pairs appear in the original documents. For example, some marketing applications may need listing the popular terms, and business intelligence applications may need listing anomalous terms as shown in Figure 19 (a) and (b) respectively. Clicking the “vis” link in Figure 19 (a) and (b) lead to the corresponding themes to which these word pairs (e.g., “energy prices” and “oil fund”) belong.
Figure 19. Drill-Down Options From Figure 10

Figure 17 also includes a list of D3 visualizations implemented. Figure 20 shows a D3 network visualization for all the data sources; their connections among the nodes are computed based on the correlations from the lexical links in Figure 17. The node connections represent all the correlations: thicker (thinner) connections indicate higher (lower) correlations. The clusters are generated based on the correlations. Figure 21 shows a D3 correlation matrix view of all the data sources. Figure 22 shows a D3 time-series bubble chart, which depicts the changing of the themes over time.
Figure 20. D3 Network

Figure 21. D3 Correlation Matrix
Task 2

We are also exploring how to use LLA jointly with other business intelligence tools, especially Big Data Architecture and Analytics (BDAA) tools:

- Deep learning, machine vision, large-scale object identification across heterogeneous data sources. One important trend in Big Data is Deep Learning, including unsupervised machine learning techniques (e.g., neural networks) for recognizing objects of interest from Big Data [9], for instance, sparse coding (Olshausen & Field, 1996) and self-taught learning (Raina et al., 2007). The self-taught learning approximates the input for unlabeled objects as a succinct, higher-level feature representation of sparse linear combination of the bases. It uses the Expectation and Maximization (EM) method to iteratively learn coefficients and bases. Deep Learning links machine vision and text analysis smartly. For example, text analysis Latent Dirichlet Analysis (LDA) is a sparse coding where a bag of words is used as the sparsely coded features for text (Olshausen & Field, 1996). Our methods Lexical Link Analysis (LLA), System-Self-Awareness (SSA), and Collaborative Learning Agents (CLA) can be viewed as unsupervised learning or Deep Learning for pattern recognition, anomaly detection, and data fusion.

A recursive data fusion methodology leveraging LLA, SSA, and CLA can be employed as follows:

- An agent \( j \) represents a sensor, operates on its own like a decentralized data fusion; however, it does not communicate with all other sensors but only with the ones that are its peers. A peer list can be specified by the agent.

- An agent \( j \) includes a learning engine CLA that collects, analyzes from its domain specific data knowledge base \( b(t,j) \)—for example, \( b(t,j) \) may represent the statistics for bi-gram feature pairs (word pairs) computed from LLA.

- An agent \( j \) also includes a fusion engine SSA with two algorithms SSA1 and SSA2 that can be customized externally. SSA1 integrates the local
knowledge base $b(t,j)$ to the total knowledge base $B(t,j)$ that can be passed along to its peers and used globally in the recursion in Figure 6. SSA2 assesses the total value of the agent $j$ by separating the total knowledge base into the categories of patterns, emerging and anomalous themes based on the total knowledge base $B(t,j)$ and generates a total value $V(t,j)$ as follows:

- **Step 1:** $B(t,j) = SSA1(B(t-1, p(j)), b(t,j));$
- **Step 2:** $V(t,j) = SSA2(B(t,j))$

where $p(j)$ represents the peer list of agent $j$.

- The total value $V(t,j)$ is used in the global sorting and ranking of relevant information. In this recursive data fusion, the knowledge bases and total values are completely data-driven and automatically discovered from the data. Each agent has the exact same code of LLA, SSA, and CLA, yet has its own data apart from other agents. This agent work has the advantages of both decentralized and distributed data fusion. It performs learning and fusion simultaneously and in parallel. Meanwhile, it categorizes the patterns and anomalous information.

- **Spark (2015):** Map/Reduce is an analytic programming paradigm for Big Data. It consists of two tasks: (1) the “Map” task, where an input dataset is converted into key/value pairs; and (2) the “Reduce” task, where outputs of the “Map” task are combined to a reduced key-value pairs. Apache Spark could replace Map/Reduce for its speed and in-memory computation.

- **Bayesian Networks with R and Hadoop (Mendelevitch, 2015):** It is a data-driven learning of conditional probability or structure learning. It is a supervised learning method but best for Big Data with low dimensions. It is an approximate inference good for Big Data and Hadoop implementation.

We have also met the acquisition professionals and discussed how BDAA can be applied to the DoD acquisition process; the following is a summary of the findings:

1. In the current acquisition process, a small delay or anomaly in a contract negotiation process can have a huge impact in its performance and can therefore cost the government a lot of money downstream.
2. It will be very useful to apply BDAA such as LLA for pattern recognition and anomaly detection for these kind of problems and make early warnings and predictions to prevent the downstream risks.
3. The Big Acquisition Data might include programs’ cost/EUM, SAR, DIMIR, tech data, people data from DMBC, even outside economic environment data if the access is possible.
4. The causes of the deviations from the normal behaviors for the programs/contracts might be modeled using physics (e.g., fluid dynamics theories).
5. LLA’s network perspectives, social plays among the nodes and the System Self-Awareness (SSA) theory may be used to lay out the academic vigor for the business processes, for example, answering the following questions:
   - Are some nodes drawn towards some other nodes because the other nodes are more powerful?
   - Is the preferential attachment growth pattern or expertise growth pattern can be used here?
• How are the forces of the nodes modeled and mapped into the social network settings and actual business processes?

Conclusion

In this paper, we show improved LLA analysis reports and visualizations generated automatically using multiple categories of data sources. These reports and visualizations reveal that there are data correlations and gaps. LLA is able to discover in detail where the gaps and inconsistencies of the data across multiple data sources reside, which, in turn, can lead to the identification of future specific and productive directions for further examination regarding why gaps occur and where they exist. It is a data-driven method for pattern recognition, anomaly detection, and data fusion. It shares indexes, not data, feasible for parallel and distributed processing, adaptive to Big Data Architecture and Analytics and needs for Big Acquisition Data.

References


## Additional Papers

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evaluating Intellectual Property and Data Rights in Competition Source Selections—Leveraging the “Assertions Process” to a New Level to Foster Open Systems Architecture</strong></td>
<td>Eugene J. Pickarz, Jr., National Reconnaissance Office</td>
</tr>
</tbody>
</table>
| **Orchestrating the Development of a Complex System of Systems: Systems Engineering Tools and Methodologies** | Carly Jackson, SPAWAR Systems Center Pacific  
Richard Volkert, SPAWAR Systems Center Pacific  
George Saroch, PMS 420 |
Abstract

Competition is repeatedly cited in the acquisition world as a powerful tool, if not the most powerful tool, to ensure taxpayers get the most value for their tax dollars. A viable competition package (Request for Proposal [RFP], Statement of Work [SOW], Performance Work Statement [PWS], etc.) is not possible without having adequate technical data, computer software, and computer software documentation to provide to potential competitors to enable them to develop or evolve a system or support the solution needed. This paper first presents the acquisition professional with the knowledge to more effectively evaluate intellectual property in source selections to ensure the Government gets the intellectual property rights it needs to procure, support, and sustain systems the warfighter and others need; second, provides a structure and process to get these “rights” identified on contract while providing transparency for them throughout the period of performance; and, finally, presents a different way to look at the “necessary” rights when viewed from an open system architecture perspective.

Introduction

Better Buying Power (BBP) 2.0 challenged the Department of Defense (DoD) to “do more without more.” One focus area was to “promote effective competition” (Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics [USD(AT&L)], 2013, p. 17). Competition is repeatedly cited in the acquisition world as a powerful tool, if not the most powerful tool, to ensure the taxpayer gets the most value for their tax dollars. It is very difficult, if not impossible, however, to develop a viable competition package (Request for Proposal [RFP], Statement of Work [SOW], Performance Work Statement [PWS], etc.) without having adequate technical data, computer software, and computer software documentation to provide to potential competitors to enable them to develop or evolve the system or support the solution needed. To this end, delivering the appropriate volume and content of technical data and computer software that are necessary to compete, support, and sustain weapon systems and their support infrastructures is critical.

Promoting effective competition was also framed in the context of using Open Systems Architecture (OSA) approaches and managing technical data rights to foster those
architectures (OUSD[AT&L], 2013, p. 18). Put quite simply, you can’t develop and maintain open architectures without access to the technical data and software they rely upon or at least utilize to some extent. BBP 3.0 does not abandon the progress of the two previous BBP releases. Rather, it continues the focus on continuous improvement with a new emphasis on initiatives that encourage innovation and promote technical excellence with the overarching goal of ensuring that the United States’ military has the dominant capabilities to meet future national security requirements. (OUSD[AT&L], 2014, p. 2)

OSA continues to be a BBP focus to stimulate technology insertion to keep pace with technology innovations and enable the design agility needed to keep ahead of our adversaries. We simply cannot effectively “refresh” our designs without the tools to foster these refresh cycles. The modularity of OSA not only stimulates innovation, but fosters competition as well from new entrants to the market from which to leverage not only commercial technology but new designs as well.

This paper is not a “how-to guide” to implement OSA from a technical perspective. It does, however, provide an approach to aid acquisition professionals in structuring RFPs, evaluating them, and making best value award decisions in competitive acquisitions. In other words, how to get OSA on contract more effectively. What is unique in this approach is that it fosters significantly improved management and insight of technical data rights and computer software toward an end goal of implementing an open systems approach both for the instant contract and those that follow. This approach results in the program managers and their acquisition teams knowing exactly what intellectual property (IP), which includes not only non-commercial data rights but also commercial software, commercial technical data, and patented inventions, are incorporated into a contractor’s technical solution and how any restrictions impact the final deliverables from a future support, sustainment, and competition perspective.

Background

The two primary parts within acquisition regulations discussed herein are the Federal Acquisition Regulation (FAR) Part 15 (Contracting by Negotiation), specifically, Subpart 15.1 Source Selection Processes and Techniques; and DoD FAR Supplement (DFARS) Part 227 (Patents, Data, and Copyrights), specifically, Subpart 227.71 (Rights in Technical Data) and Subpart 227.72 (Rights in Computer Software and Computer Software Documentation). It is where these two parts of the acquisition regulations intersect that we need to leverage to ensure the Government communicates what it needs with respect to intellectual property (IP). Getting the IP the Government needs is not an “option,” as the DoDI 5000.2 clearly levies this responsibility to program managers where it states,

Program management must establish and maintain an IP Strategy to identify and manage the full spectrum of IP and related issues (e.g., technical data and computer software deliverables, patented technologies, and appropriate

1 The essence of Open Systems Architecture (OSA) is organized decomposition, using carefully defined execution boundaries, layered onto a framework of software and hardware shared services and a vibrant business model that facilitates competition. For a full description, see (DoD, 2013, p. iii).
This “IP Strategy” can only be effectively executed when the Government knows where IP is embedded within its components, items, and processes. Program managers and their acquisition team need to be in front of the IP challenge at the beginning of acquisitions during the RFP phase. If an IP Strategy and its related issues related to missing data, computer software, and the necessary rights and licenses to use them is implemented well into a program’s schedule, it is too late to capture the savings possible through competition. Before you build an RFP, you have to first have a Source Selection Plan (SSP) that you must follow toward contract award. A brief discussion of where in the source selection process IP can be better communicated and evaluated is useful to provide context of the recommended solutions presented herein.

The Source Selection Process and IP Focus Areas

A top-level source selection process is shown in Figure 1. This figure does not attempt to capture every potential step and process (for example, conducting clarifications or awarding without discussions). It serves only to highlight where this paper identifies the impacts evaluating IP has on the overall competitive proposal/source selection process. The light shaded boxes reflect the key areas this paper will elaborate on. The SSP, the importance of evaluating and scoring/rating IP, communicating what the Government wants through the RFP, evaluating proposals, and selecting the best value offer using IP as a decision element are important to grasp the real utility of leveraging IP in competitive source selections.

- Establishing a Sound IP Strategy in Source Selections and Communicating It Clearly to Industry Is Critical

An SSP is required for all best-value, negotiated, competitive acquisitions under FAR Part 15 (OUSD[AT&L], 2011). It is within the SSP where IP can be identified as part of the evaluation criteria as either a factor or sub-factor. The more importantly IP is weighted within the total set of criteria will directly determine how much attention offerors pay to it with respect to winning the competition. The RFP must be developed to align exactly with the SSP with respect to process and the criteria to evaluate the offerors’ proposals. The RFP must also clearly communicate (through Section L, Instructions to Offerors) how to structure and present their proposal with respect to the criteria by which it will be evaluated.
Evaluation Criteria—Structure With Caution

The government has wide latitude with which to establish its requirements and needs and exercise judgment when evaluating offerors’ proposals. The General Accountability Office, in adjudicating hundreds of protests, has consistently opined that in reviewing a protest against an agency’s evaluation of proposals, it will not substitute its (or the protestor’s) judgment for that of the agency; rather, it will examine the record to determine whether the agency’s judgments were reasonable and consistent with the stated evaluation criteria and applicable procurement statutes and regulations. The evaluation of proposals is therefore a matter within an agency’s broad discretion, since the agency (not Industry) is responsible for defining its needs and the best method for accommodating them. What this means with respect to establishing IP as an evaluation criteria is that it is completely acceptable to do so. Just because an offeror is unhappy with how the IP (delivered with its solution to the Government’s requirements) was scored, does not, in and of itself, establish that the Government acted unreasonably. The Government is simply determining that the solution (with the related IP) did not represent the best value to the Government. All that being said, there are still some fundamental pitfalls that can derail an otherwise sound IP strategy.

There are some limitations with respect to IP that must be recognized and respected. The Government cannot “force” a relinquishment of rights to data (and computer software) that was independently developed at private expense. This restriction is well founded in the U.S. Code and the DFARS. That doesn’t, however, preclude the Government from identifying its minimum needs for IP and evaluating the impacts restrictive IP elements (data and computer software) have on the best value determination. There are two basic ways to evaluate IP in the competitive proposal process, scoring IP as a criteria (factor or subfactor) or as an overall IP “Risk Assessment.”

Aligning Evaluation Criteria With Ratings and/or Risk Assessments

When establishing evaluation criteria with their respective factors and subfactors, the Government must communicate not only how ratings/scores will be assigned, but also when the various standards have been met. There is great latitude with how to establish scoring methodologies, from numerical, algebraic, narrative, to adjectival. Since IP is very complex to identify and evaluate, adjectival and narrative have the most merit. An example of a previously used adjectival rating scale can be found in Table 1.

---

2 See GAO Protest Decision, B-406505; B-406405.2, dated May 21, 2012.
3 See GAO Protest Decision, B-406505; B-406405.2, dated May 21, 2012.
4 See 10 U.S.C 2320(F) and DFARS 227.7103-1(c), 227.7203-1(c).
Table 1. Technical-Management Rating Scale*

<table>
<thead>
<tr>
<th>Color</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>Exceptional. Offeror’s proposal demonstrates an exceptional understanding of the requirements, and the approach is of superior quality. Two or more significant strengths exist. There are no major or significant weaknesses, and no deficiencies exist.</td>
</tr>
<tr>
<td>Teal</td>
<td>Good. Offeror’s proposal demonstrates a good understanding of the requirements, and the approach is of good quality. Strengths clearly outbalance any weaknesses that exist. There are no significant weaknesses, and no deficiencies exist.</td>
</tr>
<tr>
<td>Green</td>
<td>Satisfactory. Offeror’s proposal demonstrates an acceptable understanding of the requirements, and the approach is of satisfactory quality. There may be strengths or weaknesses and the strengths balance the weaknesses. No deficiencies exist.</td>
</tr>
<tr>
<td>Yellow</td>
<td>Marginal. Offeror’s proposal demonstrates a marginal understanding of the requirements, and the approach is of poor quality. There may be significant weaknesses or deficiencies. Weaknesses outbalance strengths.</td>
</tr>
<tr>
<td>Red</td>
<td>Unsatisfactory. Offeror’s proposal does not demonstrate an understanding of the requirements, and the approach is of unacceptable quality. There are significant weaknesses or deficiencies. Weaknesses negate strengths.</td>
</tr>
</tbody>
</table>

* In this particular example, both Technical and Management factors were scored with the same adjectival rating scale.

One of the subfactors under the Technical-Management Factor was Innovation Approach (the highest of three total subfactors). Under this subfactor, the Government evaluated the offerors’ ability to identify and apply innovative methods of producing domain understanding and domain knowledge from multi-source, multi-dimensional data. The Government also evaluated the offerors’ ability to identify and apply innovative methods of automation human-computer interaction, data uncertainty management, and data pedigree maintenance. Lastly, the Government evaluated the offerors’ ability to minimize technology transition costs. While it’s not important for the reader to understand the technical nuances of this technical factor, it is important to focus on the last sentence of the factor from an IP perspective. This is because unless the Government has the appropriate rights and licenses to the IP necessary to execute the offeror’s technical solution, transitioning the technology developed and deployed across the Government’s organization will be cost prohibitive and potentially lead to a long-term sole source acquisition situation.

To ensure that IP independently developed at private expense (as discussed earlier) is not a “condition of offer” or that the solicitation “forces” a relinquishment of the same, the “standards” by which the subfactors will be evaluated against must clearly convey this. To this end, the standards associated with the above Innovative Approach subfactor stated that the standard would be met when the Offeror described the extent to which the rights in Technical Data (TD), Computer Software (CS), and Computer Software Documentation (CSD), and inventions/patents offered to the Government ensured unimpeded, innovative, and cost-effective production, operation, maintenance, and
upgrade of the [System Name] processing prototypes throughout its lifecycle; allow for open and competitive procurement of [System Name] enhancements; and permit the transfer of [System Name] TD, CSD, and CS to other systems or platforms.

Note the power of this one standard. In it, nine best value tradeoffs can be identified that are directly attributable to IP rights and licenses (unimpeded, innovative, cost effective, production, operations, maintenance, upgrade, future competitions, transfer to other systems or platforms).

The standard went on to ensure that proposals will not be rated less that SATISFACTORY on this standard solely because an offeror does not offer a price for all items delivered with Government Purpose Rights. However, rating on this factor for proposals to deliver TD, CSD, and CS with less than the minimum rights specified for the Government by applicable statute (10 U.S.C. 2320) and regulation DFARS 252.227-7013, 252.227-7014, and 252.227-7015 may be negatively impacted. For non-commercial acquisitions, these rights include Unlimited Rights in TD, CS, and CSD as specified in DFARS 252.227-7013 & 252.227-7014, Limited Rights in TD as specified in DFARS 252.227-7013, and Restricted Rights in CS as specified in DFARS 252.227-7014. The minimum rights considered for TD associated with commercial item acquisitions are specified in DFARS 252.227-7015. For commercial CS acquisitions, evaluation of the offered license rights will assess the licenses customarily provided to the public with respect to their consistency with Federal procurement law and satisfaction of Government user needs as set forth in the solicitation.

The key to having enough insight into the offeror’s proposal regarding the IP strategy reflected in the subfactor and its related standard is to “map” the IP within the proposal. This will be discussed later on. An alternative to “scoring” IP is to evaluate IP from an overall “Risk” perspective. To this end, an IP Risk Evaluation example is presented next.

To simplify the evaluation of IP in a source selection, some acquisition teams have chosen to assess overall IP “Risk” as reflected in an offeror’s proposed technical solution. As example of this was where the Government evaluated Intellectual Property Risk as the extent to which the Intellectual Property in technical data, computer software and computer software documentation and inventions/patents offered to the Government will:

- Ensure unimpeded, innovative and cost effective production, operation, maintenance, and upgrade of the capability/service throughout the [System Name] life cycle
- Allow for open and competitive procurement of enhancements; and will permit the transfer of technical data of non-proprietary object and code and source code to other contractors for use on other systems or platforms. (DoD, 2013)

This example used a Risk Rating table as shown in Table 2.
### Table 2. Intellectual Property Risk

<table>
<thead>
<tr>
<th>Risk Rating</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High probability for schedule disruption, cost growth, and/or performance impact. Extensive use of highly restrictive commercial or proprietary technical data, computer software, and computer software documentation, offered to the Government will definitely result in impeded, unimaginative, and cost prohibited production, operation, maintenance, and upgrade of the capability/service throughout the [System Name] life cycle, will impede open and competitive procurement of enhancements, and inhibit the transfer of the non-proprietary object code and source code to other contractors for use on other systems or platforms. Mitigation actions for commercial technical data computer software, and computer software documentation terms and conditions that are contrary to Federal law or are inconsistent with all requirements of this RFP are not identified. Risk may be unacceptable even with special contractor emphasis and close Government monitoring. Program success is jeopardized.</td>
</tr>
<tr>
<td>Moderate</td>
<td>Moderate probability for schedule disruption, cost growth, and/or performance impact. Less than Government Purpose Rights in technical data, computer software, and computer software documentation, offered to the Government may impede, innovative, and cost effective production, operation, maintenance, and upgrade of the capability/service throughout the [System Name] life cycle, may impede open and competitive procurement of enhancements, and may impede the transfer of the non-proprietary object code and source code to other contractors for use on other systems or platforms. Commercial technical data computer software, and computer software documentation terms and conditions may be contrary to Federal law, cannot be amended, and may lead to schedule disruption, cost growth, and/or performance impact. Mitigation actions for these commercial technical data, computer software and computer software documentation terms and conditions that are contrary to Federal Law or are inconsistent with all requirements of this RFP have not been adequately identified. Special contractor emphasis and close Government monitoring will probably overcome difficulties without significantly impacting program success.</td>
</tr>
<tr>
<td>Low</td>
<td>Low probability for schedule disruption, cost growth, and/or performance impact. Government Purpose Rights or Unlimited Rights in technical data, computer software, and computer software documentation, offered to the Government will definitely ensure unimpeded, innovative, and cost effective production, operation, maintenance, and upgrade of the capability/service throughout the [Program Name] life cycle, allow for open and competitive procurement of enhancements, and permit the transfer of the non-proprietary object code and source code to other contractors for use on other systems or platforms. Normal contractor effort and routine Government monitoring should overcome any difficulties. Mitigation actions have been adequately identified for commercial terms and conditions that are contrary to Federal law or are inconsistent with all requirements of this RFP. Program success is likely.</td>
</tr>
</tbody>
</table>

The "risk evaluation" approach implemented by Table 2 provides for assessing the impact of IP on the overall proposed solution across all the technical areas, vice a specific factor or subfactor as presented earlier. This gives the evaluation team even more flexibility and is actually easier to document in the IP evaluation. While both approaches, the "factor" approach and the "risk" approach, have great merit, they both require adequate clarity with respect to the identification of the IP throughout the offeror's proposal. This is facilitated by the standard "assertions process" required in DFARS 252.227-7017 and standard Section K, Representations and Certifications, Provisions, but the methodology presented herein takes these longstanding processes to a much higher level.

**Evaluating Initial Proposals**

Figuring out where IP is buried within a contractor's proposal, or more importantly, within the proposed solution, is not easy. This is because the primary enabling clauses rely upon Section K, which normally brings in the Assertions Clause, 252.227-7017 (the -7017 clause), the Prior Delivery Clause, 252.227-7028 (the -7028 clause), and the required FAR assertions pursuant to the necessary Patent clauses in the contract when applicable.
Unfortunately, even though the assertions become a part of the resulting contract, many elements of Section K are long forgotten after contract award. The net result is that IP is many times a proposal element that is overlooked, and it may come back to haunt the Government when the “impact” of the assertions become apparent upon delivery or earlier during contract performance. What is needed is a better methodology and form with which to identify and evaluate IP during the proposal evaluation process. One proven method is to leverage the assertions process.

**Leveraging the “Assertions Process” to Expand Clarity and Purpose**

*The Assertions Process*

IP can be some of a company’s most valuable assets, the relinquishment of which can significantly impact not only their profitability, but their long-term survival as well. As a result, it is in their best interests to protect them to the maximum extent possible. While only one small component of data rights management, the “assertion step” is important to understand both pre and post award as there are different standards and responsibilities tied to each. Unfortunately, the assertions required by the -7017 clause leaves a lot of uncertainty with respect to just “where” the restricted elements reside in the technical solution or services provided. The “Intellectual Property Attachment” methodology provided herein represents a best practice that “maps” the contract line items (CLINS), the Contract Data Requirements List (CDRL) items, the minimum data rights the Government has determined necessary for each deliverable, the Statement Of Work (SOW)/Performance Work Statement (PWS), the Data Rights that will be delivered; and other IP (patented inventions), all in one contract attachment with seven tables that live throughout the life of the contract. This approach facilitates efficient and thorough evaluation of IP both for initial proposals and final proposal revisions. It also establishes an additional vantage point from which to eliminate weak proposals from the competitive range and to establish another element of the “responsiveness” determination of proposals.

Rather than attempt to explain all the nuances and entitlements of the various categories of data rights, commercial technical data and software terms and conditions and patents/inventions, which are beyond the scope of this paper, the important takeaway is that the acquisition professional must clearly understand the nature and content of the technical data and computer software (both commercial and non-commercial) they identify as required to meet their minimum needs to execute their particular contract/program.

What may not be so obvious to the acquisition professional is that assertions are a critical precursor to being able to mark any deliverable containing technical data or computer software with any restrictive marking, post award. In other words, if a deliverable contains such non-commercial intellectual property, identifying the items, the basis for the restrictive marking and what restrictive category is applicable is required before delivering with a restrictive marking affixed to the specific data items. The DFARS requires these assertions be furnished to the Government and identified in “an attachment” to the contract prior to the delivery of any data with restrictive data (DFARS 227.227-7013(e)(2)). The DFARS goes on

---

5 Managing inventions and the patents that register them is not a primary focus of this paper due to the complexities of this topic and page limitations. The identification of them is however important and is presented later.
to cite the -7017 clause as the provision to facilitate this identification process, or for our purpose here the assertions necessary as a precursor to affixing restrictive markings on deliverables (DFARS 227-7103-3(b)). This Assertions List then becomes attached to the contract (DFARS 227-7103-10(a)(3)).

In like fashion, the DFARS requires the identification of computer software and computer software documentation to be furnished with restrictions prior to delivery (DFARS 252.227-7014(e)(2)). As with Technical Data, the -7017 clause is used again to facilitate the same due diligence actions by the contractor discussed earlier. It bears repeating again that unless a restriction is asserted, no restrictive markings may be affixed to the final software. (This is normally done via “code headers” within the software itself and the marking of the physical documentation of the software.) Both Government and contractor alike should take extreme care during the software acceptance process to ensure that non-commercial computer software is scanned to identify any internal restrictive markings as they can coexist with a transmittal letter that alludes to something else. Once incongruent markings are identified, the corrective actions may be invoked as set forth in both the -7013 clause and the -7014 clause.

Before we get to the details of the Assertion List itself developed pursuant to the -7017 clause, the causal link between assertion and delivery is useful to revisit. If you read both the -7013 and -7014 clauses carefully you will note that the activity of delivery is woven throughout. Thus the action of delivery is required to empower the Government to assert its data rights on the non-commercial Technical Data or Computer Software in question. As explained previously, any restrictions must be “asserted” prior to any such delivery. But before any such assertions may be made, the specific technical data or computer software must be “identified” as required for delivery. This is an important sequence of events that must take place to effectively manage data and the protection thereof. In other words, no assertion, no restrictive marking authorized if you are the contractor. But if you are the Government, beware, because without a requirement for “delivery” the contractor is not bound to identify or assert any restrictions. Only if delivery is later called for (via deferred ordering) or identified as a post award assertion (which has more strict limitations than pre-award assertions) will the identification and the restrictions be brought to light.6

Let’s now turn to one of the key elements of the discussion, namely, the Assertions List, or more importantly the -7017 clause elements that lead to the “List” or “Attachment” itself. This is the traditional methodology (combined with attaching commercial software licenses to the contract and citing patent royalty information in Section K, as discussed earlier).

Since this is so critical to the discussion here, the elements of the clause are provided in Table 3.

---

6 See DFARS 252.227-7013(e)(3), 252.227-7014(e)(3), and 252.227-7018(e)(3).
Table 3. -7107 Clause Elements (DFARS 252.227-7017(e)(3))

The Offeror asserts for itself or the persons identified below, that the Government’s rights to use, release, or disclose the following technical data or computer software should be restricted:

<table>
<thead>
<tr>
<th>Technical Data or Computer Software to be Furnished With Restrictions*</th>
<th>Basis for Assertion**</th>
<th>Asserted Rights Category***</th>
<th>Name of Person Asserting Restrictions****</th>
</tr>
</thead>
<tbody>
<tr>
<td>(LIST)*****</td>
<td>(LIST)</td>
<td>(LIST)</td>
<td>(LIST)</td>
</tr>
</tbody>
</table>

*For technical data (other than computer software documentation) pertaining to items, components, or processes developed at private expense, identify both the deliverable technical data and each such item, component, or process. For computer software or computer software documentation, identify the software or documentation.

**Generally, development at private expense, either exclusively or partially, is the only basis for asserting restrictions. For technical data, other than computer software documentation, development refers to development of the item, component, or process to which the data pertain. The Government’s rights in computer software documentation generally may not be restricted. For computer software, development refers to the software. Indicate whether development was accomplished exclusively or partially at private expense. If development was not accomplished at private expense, or for computer software documentation, enter the specific basis for asserting restrictions.

***Enter asserted rights category (e.g., Government purpose license rights from a prior contract, rights in SBIR data generated under another contract, limited, restricted, or Government purpose rights under this or a prior contract, or specially negotiated licenses).

****Corporation, individual, or other person, as appropriate.

*****Enter “none” when all data or software will be submitted without restrictions.

Date
Printed Name and Title
Signature

Columns two through four are the easiest to deal with. Column four is very straightforward: who is the right person to sign off. Column three is fairly simple as well: it’s either Restricted Rights (for software), Limited Rights (for technical data, (SBIR Rights), or Government Purpose Rights (where mixed funds are/were used). Column two, the basis column, is pretty straightforward as well, and it usually reads “Independently Developed at Private Expense” or “Jointly Developed with Contractor and Governments funds.” If you look at the information sought in column one, however, it may be interpreted in some instances ambiguously. Just what is required to “identify the technical data, computer software or computer software documentation”? An ambiguous assertion example could be “All XYZ software utilized in the ABC assembly.” This “notional” top level data description is extremely

---

7 Note that the -7013, -7014, and -7018 clauses all have an identical table with some of the instruction language that is to be used for post award assertions.
problematic for two main reasons. First, the Government has no clue “what” software is being restricted (assuming, of course, Column three would indicate “Restricted Rights”), and second, the Government really doesn’t know clearly “what” software it really should be protecting with that level of restriction, where within the system or architecture design the restricted software resides, nor what it is really getting for its money. True, it is easy to simply say “all,” but is it fair and accurate? Most would agree, it’s not. You can’t really determine what you “need” to field, support, and sustain a system without knowing what you “have” to begin with (Pickarz, September 2012). In this instance, you just don’t know and most importantly, you have no baseline at contract award from which to later determine what changes the Government has funded and may have unlimited rights to. A huge entitlement may be lost from simply not paying attention to the assertions contained in the Assertions List. The solution to this dilemma is actually quite simple. Namely, make the instructions unequivocally clear. A formal deviation to the clause is probably not a timely solution. A better solution is clearer instructions to the contractor in the solicitation in Section L (Instructions to Offerors) with a resultant attachment to the contract that documents the technical data and computer software and their respective rights to be given to the Government. It is much better to articulate just what you expect the contractor to deliver in their proposal rather than have them guess. For the example earlier, the software version(s), and/or dates should be given to clearly identify just what will be restricted upon delivery. Even better, if you make clarity of the assertions a condition of offer, contractors will always comply or possibly lose the award. Let me be clear, however, as the DFARS deals with this very situation where it states,

If an offeror fails to submit the attachment or fails to complete the attachment in accordance with the requirements of the solicitation provision, such failure shall constitute a minor informality. Provide offerors an opportunity to remedy a minor informality in accordance with the procedures at FAR 14.405 or 15.607. An offeror’s failure to correct the informality within the time prescribed by the contracting officer shall render the offer ineligible for award. (DFARS 227.7103-10(a)(1))

Note that while clarity would be considered a “minor informality,” failure to correct this shall render the offer ineligible for award. Another key point is that a minor informality could be resolved as a “minor error” pursuant to a “clarification” vice a “discussion” point, thereby preserving the ability to award without discussions should this be provided for in the solicitation (FAR 15.306(a)(2)). At the end of the day, additional emphasis in the instructions for completing the assertions goes a long way to enable the Government to later assert the rights it has paid for.

The Intellectual Property Attachment—Mapping Critical IP Artifacts

Non-commercial technical data and computer software assertions are really only part of the intellectual property portfolio as there are numerous commercial technical data and computer software artifacts, and in many cases previously developed inventions, that are relevant to Government contracts. The answer to the question, “What do I have?” is important not only at contract award but throughout contract performance as the

---

8 See DFARS 252.227-7013(b)(1), 252.227-7014(b)(1).
deliverables from one contract provide the building blocks for another contracts and their programs/projects. All this assertion information can be captured in one place both to evaluate the proposal and then continue throughout contract performance as a living document. This is accomplished by adding an “Intellectual Property Volume” to your solicitation and the resulting “Intellectual Property Attachment” to the awarded contract.

This original idea was first promulgated by Space and Missile Center (SMC) in Los Angeles and presented in the SMC Office of the Staff Judge Advocate Guide *Acquiring and Enforcing the Government’s Rights in Technical Data and Computer Software Under Department of Defense Contracts: A Practical Handbook for Acquisition Professionals* (Space and Missile Center [SMC] Office of the Staff Judge Advocate, March 2014). Now in its Sixth edition, this somewhat daunting document may seem to be a bit difficult to review at first, but searching for the “Data Rights Attachment” will get you to most of the components discussed herein. For the purposes of this paper, I will take SMC/JA’s approach and expand it to provide a comprehensive “Volume” to the proposal that lays out not only the Data Rights attributable to the effort but other areas of intellectual property as well. To do this, an Intellectual Property Volume is required from offerors. This volume would be structured as follows:

Volume “X”—Intellectual Property

- Table 1—Data Rights Summary: Non-Commercial Technical Data and Computer Software & Computer Software Documentation
- Table 2—Commercial Technical Data and Computer Software & Computer Software Documentation
- Table 3—Assertions List: Non-Commercial Technical Data, Computer Software, and Computer Software Documentation
- Table 4—Specifically Negotiated Licenses (Special Licenses to Non-Commercial Technical Data and Computer Software)
- Table 5—Rights in Background Inventions
- Table 6—Third Party Patent Rights and Royalties

It helps to visualize the Intellectual Property Volume approach so the following notional tables with example deliverable technical data and computer software deliverables are provided. The various elements of the tables and their mapping functions will be discussed.

---

9 While this paper focuses on “data rights,” Tables 6 and 7 are provided and briefly discussed to add the listing of any relevant inventions (Patents) used in the contractor’s proposed solution. This incorporation then provides a comprehensive IP attachment to the contract.
Table 4. Data Rights Summary
Non-Commercial Technical Data and Computer Software and Computer Software Documentation

<table>
<thead>
<tr>
<th>Col. 1</th>
<th>Col. 2</th>
<th>Col. 3</th>
<th>Col. 4</th>
<th>Col. 5</th>
<th>Col. 6</th>
<th>Col. 7</th>
<th>Col. 8</th>
<th>Col. 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0002</td>
<td>A0001</td>
<td>Computer Software</td>
<td>3.1.5</td>
<td>Government Purpose Rights</td>
<td>Offerer To Complete</td>
<td>Offerer To Complete</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>0002</td>
<td>A0002</td>
<td>User Manuals</td>
<td>3.1.4</td>
<td>Unlimited Rights</td>
<td>N/A</td>
<td>Unlimited Rights</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>0002</td>
<td>A0003</td>
<td>Contract Funds Status Report</td>
<td>3.1.6</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>0002</td>
<td>A000n</td>
<td>Title</td>
<td>Para Nr. &quot;n&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. This column is normally used for competitive Requests for Proposals (RFPs). Here the offeror(s) indicate what their usual entitlements are for the respective data deliverable. For example if the data was independently developed at private expense, Restricted Rights (Software) and Limited Rights (Technical Data) are usually appropriate. It is important to note that contractors cannot be forced to relinquish their rights related to independently developed data and this should be clearly stated in the RFP. They are, however, able to propose higher levels of data rights in order to be more competitive in the marketplace which is facilitated by Column 7 in the table above. In both instances, the offeror fills in these blocks of the table.

2. In this instance, the Government has filled in the table because it has determined that the User Manuals constitute Operation, Maintenance, Installation, and Training (OMIT) data and should be delivered with Unlimited Rights in accordance with 10 U.S.C. 428(3)(b)(ii). In other words, the offerors don’t get to restrict the data, nor do they get to propose their “usual rights.” If there is a disagreement, the offeror will note it in Column 8, and perhaps the RFP can provide for Data Rights Options to price out additional entitlements the Government may purchase if the Government’s identified minimum rights cannot be offered in the established line item prices.

3. Column 8 affords offerors the opportunity to explain why they will not provide the required data item with the indicated Government rights. It also affords the opportunity to propose a "Special License." Special licenses are cited in Table 3.

4. The Price/Estimated Cost column is only necessary when separately pricing data items. It is important to note that when separately priced data items are required, additional pricing instructions in Section L are necessary to guide offeror.

As you can see from the Data Rights Summary table, the data rights the Government will receive are clearly “mapped” to the contract’s Contract Line Items (CLINs), Contract Data Requirements Lists (CDRLS), and the SOW/PWS. CLIN 0002 in this particular solicitation was for “Data,” and a few notable items are presented. But there’s some important nuances to take note of that reflect the true power of this approach. Note first that the Government has clearly identified what its minimum needs are for this acquisition in Column 5. Note also that the User Manuals constitute OMIT data, which entitles the Government to Unlimited Rights, so this cell in the table has been “pre-filled” to establish this entitlement. The Contract Funds Status Report (CFSR) is marked N/A. This is because the CFSR constitutes financial data that is incidental to contract administration and outside the definition of “Technical Data,” which triggers the applicability of the various rights outlined in the clauses. Finally, Column 8 provides the ability for offerors to explain why the rights proposed do not meet the Government’s minimum needs (again to preclude forcing the relinquishment of rights to independently developed technical data or computer software.) This table from the proposal will become an attachment to the contract and a “living” document (as will all the tables discussed here) to provide for adding post award assertions and afford the Government complete Intellectual Property situational awareness.

Table 5 provides the insight to any commercial technical data or computer software the contractor must deliver under the contract. This table contains nine columns.
Table 5. Commercial Technical Data, Computer Software, and Computer Software Documentation

<table>
<thead>
<tr>
<th>Col. 1</th>
<th>Col. 2</th>
<th>Col. 3</th>
<th>Col. 4</th>
<th>Col. 5</th>
<th>Col. 6</th>
<th>Col. 7</th>
<th>Col. 8</th>
<th>Col. 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIN</td>
<td>CDRL</td>
<td>SOW/PWS Paras</td>
<td>Data Item Title/Subtitle</td>
<td>Technical Data/Software Application Name</td>
<td>Vendor Name</td>
<td>License No.</td>
<td>Quantity of Licenses (If applicable)</td>
<td>Price/Estimated Cost</td>
</tr>
<tr>
<td>0002</td>
<td>A0001</td>
<td>Para Nrs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0002</td>
<td>A0002</td>
<td>Para Nrs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. The instructions for this column in Section L must clearly articulate the requirement for clarity and thoroughness in the item descriptions for the reasons mentioned earlier. It needs to be made clear to offerors that cryptic and incomplete descriptions may render their offer non-responsive to the solicitation. In addition, if firmware is to be delivered, the Data Item Title should include the CLIN Noun Description that the firmware will be embedded into.

In like fashion of the other tables, the first three columns provide for CLINs, CDRLs and SOW/PWS paragraphs that are mapped to the commercial technical data and commercial computer software that are to be delivered that is provided for in the proposal. Instructions in Section L will again guide offerors how to populate the table. The Government needs to ensure that identifying commercial software is not enough, and Open Source Software (OSS) and other openly shared software must also be identified since they are also commercial products in nature. This is because even though a software artifact may be “open,” it still has terms and conditions by which it must be shared. The commercial license terms can be problematic and the Government may have concerns regarding these commercial technical data and computer software licenses and these must be adjudicated. Some of these concerns relate to the following:

- Subsequent rights to updates, software maintenance patches, minor version changes and substitutions provided at no additional cost
- License transferability to the Government (for option exercise and CDRL/CLIN delivery)
- Disputes provisions
- Choice of law provisions
- Payment of attorney’s fees
- Automatic renewal provisions that violate the Anti-Deficiency Act
- Provisions that prohibit disclosure of license terms/conditions
- Open Source Software terms that mandate sharing and posting of changes when doing so may jeopardize national security

Of course the Government has no idea if any of these unwanted terms are embodied within the commercial licenses unless the offeror is instructed to actually provide all licenses as an addendum to the table in the IP Volume. Once provided, the Government can perform its due diligence. There have been instances where offerors have claimed that license terms cannot be provided until the licenses are executed after award and failed to provide copies of the standard licenses normally required from commercial vendors. This argument is not completely true. While it is true that the final license will reflect the actual terms and conditions agreed to, virtually every commercial software product (or standard technical data documentation) has a standard license that is at least the starting point for negotiating the final terms. These “standard” licenses must be provided to enable a thorough proposal review and to develop clarification questions, information requests, and assign strengths, weaknesses, or deficiencies. In the event terms that are not acceptable to the Government
are unable to be removed, then it is a good practice to establish an overall Intellectual Property “Risk” rating to capture the additional risk to the Government from the restrictive terms as was discussed earlier.

The Assertions List, generated in response to the -7017 clause, is provided for in Table 6 of the IP Volume and again maps the restrictions and data rights proposed to the Government’s requirements laid out in the CLINs, CDRLs, and SOW/PWS. The additional benefit this approach establishes is that the clarity needed to effectively manage the Technical Data and Computer Software is mandated as a consideration of responsiveness to the solicitation. It is important to understand the difference between Table 4, which identifies the overall data rights assigned to the various data items, to Table 6, The Assertions List. Table 4 assigns the data rights, but Table 6 identifies the specific restrictive items (if any) that are tied to the restrictions. In other words just “what” makes the deliverable Limited Rights technical data. These assertions are also required for those instances where the Government identifies Government Purpose Rights (GPR) as its minimum and the contractor proposes GPR. This is because there are still elements or activities of GPR that provide for additional due diligence on the part of the Government when sharing with third parties (additional Non-Disclosure Agreements, for example). The Assertions List would thus look similar to that shown in Table 6.

### Table 6. Assertions List—Non-Commercial Technical Data, Computer Software, and Computer Software Documentation

<table>
<thead>
<tr>
<th>Col. 1</th>
<th>Col. 2</th>
<th>Col. 3</th>
<th>Col. 4</th>
<th>Col. 5</th>
<th>Col. 6</th>
<th>Col. 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIN</td>
<td>CDRL</td>
<td>SCW/PWS</td>
<td>Technical Data or Computer Software to be Furnished with Restrictions</td>
<td>Basis for Assertion</td>
<td>Asserted Rights Category</td>
<td>Name of Person Asserting Restrictions</td>
</tr>
<tr>
<td>0002</td>
<td>A0001</td>
<td>Para Nrs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0002</td>
<td>A0002</td>
<td>Para Nrs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. The instructions for this column in Section L must clearly articulate the requirement for clarity and thoroughness in the item descriptions for the reasons mentioned earlier. It needs to be made clear to offerors that cryptic and incomplete descriptions may render their offer non-responsive to the solicitation.

The “prior delivery list,” generated in response to the -7028 clause, is provided for in Table 7 and maps the technical data and computer software that was delivered to the Government prior to the current effort (or is scheduled to be delivered on another ongoing contract). Readers should keep in mind, however, that unless there were deliveries earlier in time (or planned for the future) that would be subject to reporting in the table, the offeror will simply report “none.” Again, delivery is paramount for the successful functioning of various clauses and the rights they impose. In addition to the standard information required, the relevant CDRLs are identified, as well as all contract information from which the items were/are to be delivered that are identical or substantially similar to documents or other media that the offeror has produced for, delivered to, or is obligated to deliver to the Government under any contract or subcontract (DFARS 252.227-7028).
Table 7. Prior Delivery List for Technical Data or Computer Software

<table>
<thead>
<tr>
<th>Col. 1</th>
<th>Col. 2</th>
<th>Col. 3</th>
<th>Col. 4</th>
<th>Col. 5</th>
<th>Col. 6</th>
<th>Col. 7</th>
<th>Col. 8</th>
<th>Col. 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIN</td>
<td>CDRL</td>
<td>TDICS</td>
<td>Previously Delivered</td>
<td>Contract Nr. Under Which Data Was Most Recently Delivered or Will Be Delivered</td>
<td>Organization Name &amp; Address to Which the Data Was or Will Be Delivered</td>
<td>Delivery Date</td>
<td>Limitations on Gov't's Right to Use or Disclosure Data</td>
<td>Date Limitation(s) Expire</td>
</tr>
<tr>
<td>0002</td>
<td>A0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0002</td>
<td>A0002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The instructions for this column in Section L must clearly articulate the requirement for clarity and thoroughness in the item descriptions for the reasons mentioned earlier. It needs to be made clear to offerors that cryptic and incomplete descriptions may render their offer nonresponsive to the solicitation.

Table 8 constitutes the identification of any Special Licenses relevant to a specific CDRL data item. It is important that the scope or terms and conditions of any special license be clearly articulated in the proposal and a copy of the actual license to be executed be provided as an Addendum to Table 8 for subsequent review, evaluation by the Government, and incorporation into the contract as an attachment. A notional format for Table 8 can be found below.

Table 8. Specifically Negotiated Licenses (Special Licenses)—Non-Commercial Technical Data and Computer Software

<table>
<thead>
<tr>
<th>Col. 1</th>
<th>Col. 2</th>
<th>Col. 3</th>
<th>Col. 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIN</td>
<td>CDRL</td>
<td>SOW/PWS Paras</td>
<td>Special License Title/Number/Version</td>
</tr>
<tr>
<td>0002</td>
<td>A0001</td>
<td>Para Nrs.</td>
<td></td>
</tr>
<tr>
<td>0002</td>
<td>A0002</td>
<td>Para Nrs.</td>
<td></td>
</tr>
</tbody>
</table>

1. The instructions for this column in Section L must clearly communicate that the scope of the license must be articulated and provided with the proposal. This includes identifying who the CDRL may be released or disclosed to, the purposes of the license, and the period of time it shall be in effect. The instructions must also require attaching the Special License as an addendum to Table 8 in the IP Volume.

Table 9 provides the insight to any inventions the contractor plans to incorporate into any component, item, or process. A “background invention” is any invention, other than a subject invention, that is covered by any patent or pending patent application in which the offeror (including its sub-offerors or suppliers, or potential sub-offerors or suppliers at any tier) (1) has any right, title, or interest; and (2) proposes to incorporate into any items, components, or processes to be developed or delivered, or that will be described or disclosed in an technical data, computer software, or computer software documentation to be developed or delivered under the resulting contract (DoD, May 2013). This table contains six columns.
Table 9. Rights in Background Inventions

<table>
<thead>
<tr>
<th>CLIN</th>
<th>CDR</th>
<th>SOW/PWS Paras</th>
<th>Patent Title/Serial Nr.</th>
<th>Patent Date</th>
<th>Patent Holder Name(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0002</td>
<td>A000</td>
<td>Para Nrs.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0002</td>
<td>A000</td>
<td>Para Nrs.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. The Patent rights clauses of the solicitation provide additional guidance regarding the identification and management of inventions used on the contract. The purpose of this table is to capture the inventions/patents to be incorporated in one place as an attachment to the contract to preclude losing their identity as part of the technical baseline after the initial award is completed.10

Table 10 provides the insight to any third-party patent rights for which the contractor plans to pay royalties. This table provides information concerning these third-party patents and the amount of the royalties it will pay in order to perform under the contract. This table contains seven columns.

Table 10. Third Party Patent Rights and Royalties

<table>
<thead>
<tr>
<th>CLIN</th>
<th>CDR</th>
<th>SOW/PWS Paras</th>
<th>Patent Title/Serial Nr.</th>
<th>Patent Date</th>
<th>Patent Holder Name(s)</th>
<th>Royalties to be Paid</th>
</tr>
</thead>
<tbody>
<tr>
<td>0002</td>
<td>A000</td>
<td>Para Nrs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0002</td>
<td>A000</td>
<td>Para Nrs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. See FAR Clause 52.227-6 for specific guidance on Royalties. This language (and table) only becomes relevant if a third-party patent is included in the technical solution proposed (or incorporated post award). The purpose of this table is to capture the patent royalty to be paid in one place in an attachment to the contract to preclude losing their identity as part of the technical baseline after the initial award is completed.

**Communicating the Government’s Expectations Is Vital to Success**

The acquisition team crafting the RFP needs to pay close attention when drafting the instructions in Section L related to the Intellectual Property Volume. Explaining what is needed within the various tables ensures all offerors have a common understanding. Trying to avoid specifying the “table format” and just provide Section L language tends to give inconsistent results and lead to more clarifications and/or discussions. Some offerors will interpret the instructions differently, and the result is data rights information spread across the proposal and is a virtual “scavenger hunt” to figure out just what IP impacts there are in

10 The identification of inventions to be used in a contractors technical solution described in their proposal is normally required in Section K of the solicitations. This completed section is, however, buried in the contract file documentation at contract award and can be difficult to locate in a voluminous contract file. The table methodology presented here keeps it at the forefront of the acquisition team to manage throughout the contract period of performance.
offerors’ proposals. The best approach is a Tabular One by IP topic and then incorporate the individual tables as “Tabs,” with clear instructions to populate the tables. Finally, the Government must ensure that updates to the various IP tables in the IP Attachment are reviewed and approved prior to incorporating the technical changes reflected in the projected updates during post award performance. This is to ensure configuration management of the technical changes is carefully managed and maintained.

Open Systems Architecture (OSA)—Where Does It Fit?

Open Systems Architecture provides for designs that accommodate updated technology (data and computer software) by leveraging modular, loosely coupled and highly cohesive components within a system. A system should be designed in major “modules” where potentially proprietary data and/or computer software is “encapsulated” (i.e., segregated within the design). These modules must be “loosely coupled” whereby individual modules do not depend upon each other to enable the entire system to function.) Lastly, the modules must be highly cohesive so the module functionality works together via common standards. The system relies on open interfaces well known by all competitors to enhance future competitions as well as more effective sustainment and supportability. This approach enables even highly restrictive and even proprietary designs to be incorporated into the final system yet still enable technology insertion with new and innovative upgrades. It is only made possible, however, if the “critical” IP components are or have been delivered to the Government earlier in the system’s life cycle. Knowing where the IP is embedded within the various designs fosters this approach as well by enabling strategic decisions where to focus on “opening up” the system for more competition and technology insertion. The IP Volume discussed earlier provides for the situational awareness necessary to bring it all together. For a thorough discussion on Open Systems Architecture and multiple examples and guides, readers should review the OSA Guidebook for Program Managers (DoD, May 2013).

Government Insistence on Additional Openness and IP Rights—Is It Viable?

It is important to discern whether or not the Government, in implementing the processes and strategies presented here can be sustained and implemented when challenged by Industry. There are numerous Government Accountability Office (GAO) bid protest decisions that have upheld the Government’s decisions which supports the concepts presented herein. What is important to take away from these decisions are some key principles that when adhered to, result in new IP and OSA strategies that are executable and sustainable when challenged.

As stated earlier, if the Government establishes a plan to evaluate proposals (a source selection plan), follows the plan, consistently applies the criteria and their standards fairly to all offerors, then makes a best value decision based upon all evaluation areas (cost and non-cost), the GAO will not overturn the Government’s decision. This has been a consistent result in multiple bid protest decisions. Can the Government use Open Architecture as a criteria in source selection? Can it require offerors to clearly identify what data rights the Government will obtain with an offeror’s proposed design/solution? Finally, can the Government make a best value decision using Open Systems Architecture (OSA/OA) and the delivered data rights for the technical data and computer software artifacts? The answer to all is unequivocally, yes.
A recent bid protest concerning an Engagement Skills Training (EST) system will help illustrate.\textsuperscript{11} In this instance, the Government provided for an Open Architecture subfactor to assess the ability of the offeror’s design to “fully support, maintain, and modify the EST software and technical data throughout the program life cycle to include the legacy EST systems, weapons and scenarios.”\textsuperscript{12} The acquisition team evaluated the proposals using the scales and criteria called out in their RFP. During the evaluation the team identified several areas where the Government’s license rights were cited inconsistently in different sections of an offeror’s proposal. Because of this, the evaluation team was unsure just what rights the Government would receive. A lower “marginal” rating under the open architecture subfactor was then assigned. The ambiguity was created by the offeror in the errors it submitted related to a material aspect of the technical approach regarding open architecture.\textsuperscript{13} Specifically, the inability of the Government to share many IP artifacts of technical data and computer software. The unsuccessful offeror challenged other areas of the evaluation but these will not be recounted here for the sake of brevity. The lessons learned are important, however. First, if the Government communicates what it will evaluate, how it will evaluate, and what will be taken into consideration in the best value decision, the GAO will support the Government’s decision. Second, it’s not the Government’s job to “rewrite” the offeror’s proposals and identify each and every error and weakness identified. The FAR requires agencies conducting discussions with offerors to address, “at a minimum … deficiencies, significant weaknesses and adverse past performance information to which the offeror has not yet had an opportunity to respond” (FAR 15.306(d)(3)). The Government does not have to “spoon feed” an offeror as to each and every item that could be revised to improve and offeror’s proposal.\textsuperscript{14} Finally, data rights can be directly rated and scored in a competitive source selection to enable the Government to make a best value decision.

In another bid protest decision, the Government’s requirement was for a commercial off-the-shelf, “web-based, automated e-Recruitment solution, including all software, software documentation, implementation support, and services to support the full life cycle of an enterprise-wide hiring/recruitment system.”\textsuperscript{15} An unsuccessful offeror did not include an adequate explanation, as requested by the solicitation, of the proposal’s compliance with the solicitation’s minimum mandatory requirements concerning intellectual property/data rights.\textsuperscript{16} At issue were terms of the license whereby the agency’s data once entered into the offeror’s database became the property of the offeror. This was because a term of the license required all data be identified prior to contract start. Since the goal of the project was to manage employment and other HR data throughout the period of performance, this did not meet a material requirement of the RFP which was clearly called out in a mandatory “functional requirements matrix.” The offeror’s proposal was scored commensurately and they were eliminated from the competitive range. This protest illustrates a critical lesson relevant to our discussion here. Namely, establishing material requirements in an RFP is something that standard commercial licenses may be in conflict with. Recall the language in both the “Scoring” and “Risk” approaches to evaluating IP discussed earlier. In both,

\textsuperscript{11} See GAO Bid Protest B-410006; B-410006.2, dated October 8, 2014.
\textsuperscript{12} See GAO Bid Protest B-410006; B-410006.2, dated October 8, 2014.
\textsuperscript{13} See GAO Bid Protest B-410006; B-410006.2, dated October 8, 2014.
\textsuperscript{14} See GAO Bid Protest B-404671.2, B404671.4, dated April 8, 2011.
\textsuperscript{15} See GAO Bid Protest B-298380.4, dated June 11, 2007.
inconsistency with the “requirements of the RFP” and “satisfaction of Government user needs as set forth in the solicitation” were key discriminators. Thus, “hiding behind the commercial item veil” as it were, to claim that a standard commercial license may not be challenged regarding its terms and conditions is not sufficient to negate the basic needs of the Government for IP that effectively meet their needs. What is important is to provide an adequate license to meet the Government’s requirements called out in an RFP.

Offerors should provide their best initial proposal in response to the Government’s RFP or risk being eliminated from the competitive range. This is an important point to understand as was illustrated in a recent bid protest where an offeror failed to provide significant material data and information required by RFP In Section L. The GAO has opined previously that “an offeror has the responsibility to submit a well-written proposal, with adequately detailed information that clearly demonstrates compliance with the solicitation requirements and allows a meaningful review by the procuring agency.” The offeror in this instance admitted it failed to provide information requested by the RFP and as a result their proposal failed to demonstrate that it met the solicitation requirements and they were eliminated from the competitive range. The lesson relevant to the discussion here, specifically to the IP Summary Attachment/Volume described earlier, is that unless an offeror pays close attention to the detailed instructions for this volume they run the risk of being eliminated from the competitive range. This is especially true when IP and the associated license rights and license terms and conditions are necessary to make a best value decision that has decision criteria based on IP and/or Open Architecture.

Conclusion

The goal of this paper was three-fold: first, to present the acquisition professional with some tools to ensure the Government gets the intellectual property rights it needs to procure, support, and sustain the systems the warfighter, and others, need; second, to provide a structure and process to get these rights identified on contract while providing transparency into them throughout the period of performance and not finding out “upon delivery” what rights are really being delivered; and finally, to present a different way to look at the “necessary” rights when viewed from an open architecture perspective. This is facilitated by strategically seeking the “necessary” IP rights (based upon the Government’s minimum needs) that focus on interfaces and other artifacts to implement an OSA approach. When this approach is implemented at the onset of a contract/program, restricted and limited rights become mitigated inhibitors to technology insertion and instead become catalysts to enable more affordable support, sustainment, and cost effective systems and solutions for the Government.

References


---


Orchestrating the Development of a Complex System of Systems: Systems Engineering Tools and Methodologies

Carly Jackson—is a systems engineer with SPAWAR Systems Center (SSC) Pacific and a Distributed Surveillance Systems branch supervisor. She is currently supporting the LCS MM Program Technical Director’s office, primarily tasked with the implementation of product and process commonality across complex systems of systems, systems engineering preparations for the impending Milestone B, technology transition planning, and overall management of the Mission Package Technology Development Working Group. Jackson earned simultaneous BS and MS degrees in mechanical engineering from UCLA in 2002 and an MBA in business administration from Pepperdine University in 2007. She is Level III certified by the DAU in SPRDE-Systems Engineering and Level II certified in Program Management. [carly.jackson@navy.mil]

Rich Volkert—is the lead systems engineer supporting SSC Pacific, and the PMS 420 deputy technical director. He has over 28 years of service in the government, including 20 years as an active duty naval officer with service as a engineering duty officer and in submarines. Over 19 years of that time he has been involved in the fields of research, development, acquisition, and systems engineering. He possesses degrees in aerospace engineering and acoustical engineering and is presently enrolled in a PhD program for systems engineering. He is Level III certified by the DAU in SPRDE-Systems Engineering, Program Systems Engineering, Test and Evaluation, and Program Management. richard.volkert@navy.mil]

Introduction

Acquisition efforts are under increasing pressure to deliver systems rapidly and at lower cost while providing enhanced performance capability. This is happening in an environment where research and development budgets are decreasing across the Department of Defense (DoD), the acquisition workforce is rebuilding, and the DoD is looking to rebalance warfighting portfolios. This has been compounded by an extremely dynamic environment of fiscal change and uncertainty imposed by significant budget cuts and sequestration impacts in the recent past.

The overall trend of DoD budgets has been declining. The president’s fiscal year (FY) 2015 budget recognizes the fiscal imperative of deficit reduction, reducing projected defense budgets by about $113 billion over five years compared to levels requested in the FY2014 budget (DoD, 2014). As a percentage of Gross Domestic Product (GDP), the DoD budget has been declining from approximately 15% in the mid-1950s to around 3% today. The growth in the DoD’s costs under the Congressional Budget Office (CBO) projection of the base budget is somewhat less than the CBO’s (2014) projection of the growth of the U.S. economy, decreasing slowly from 2.8% of GDP in 2015 to 2.5% in 2025, and 2.3% in 2030.

One outcome of the fiscal pressures has been an increase in defense acquisition efforts that are focused on the rapid development and fielding of an integrated system of systems (SoS) capability. Many of these SoS leverage mature systems and sub-systems as “constituent” systems but do not contract directly with industry to deliver the integrated capability. Elements of the DoD acquisition workforce are therefore taking on increasingly technical roles and responsibilities as the lead integrators for complex SoS. It should also be noted that the acquisition workforce (AWF) has declined by more than 50% since the 1990s, while the value of Department of the Navy (DoN) contracting has increased by more than 50%. With these significant AWF losses and attendant workload increases, the DoN recognized the loss in ability to manage the technical-cost tradespace it is responsible to execute, including major weapons systems acquisition (DoN, 2010). While the DoN AWF has grown by more than 20% (DoN, 2015) since 2010, some program managers are still
challenged with staffing and resourcing an acquisition workforce with the technical depth to assume the increasingly technical roles associated with the development and fielding of an inherently government SoS capability. In many cases, the Naval Warfare Center and Systems Center’s engineering workforce is tasked to supplement the acquisition workforce with technical subject matter expertise in systems engineering, rapid prototyping, development and integration, test and evaluation, operations analysis, and other inherently government roles.

Several novel systems engineering tools and methodologies have been developed to support the lead SoS engineer (LSoSE) with assessing system performance, system readiness, and risk through incremental development of a complex SoS. This paper will review the systems engineering processes and required level of technical rigor to manage the development and integration of a complex SoS in a distributed environment. A specific implementation of a Kill Chain framework that was designed to provide the PM and LSoSEs with comprehensive insight into the technical status and major risks for a complex SoS across specific mission threads will be presented. This paper will also review some of the systems engineering tools and methodologies developed and implemented by the Littoral Combat Ships (LCS) Mission Modules (MM) Program Office (PMS 420) that support the LCS MM Kill Chain framework.

**LCS Program Overview and Background**

**Acquisition Strategy**

The Flight Zero Capability Development Document (CDD) for the Littoral Combat Ship (LCS) was approved in April 2004 and established the requirements of the LCS and three “focused” missions. Each Mission Package (MP) provides warfighting capability for one of three focused mission areas:

- Mine Countermeasures (MCM)—Detection and neutralization of mine threats
- Surface Warfare (SUW)—Maritime security missions and fleet protection from small boats and other asymmetric threats
- Anti-Submarine Warfare (ASW)—Countering shallow water diesel submarine threats

The CDD was updated in 2008 to the current Flight 0+ version which included several requirement updates, which were used to drive the design for the full “production” versions of the seaframe and MPs. MP specific requirement updates included the Net Ready Key Performance Parameter (KPP), Material Availability KPP, and others.

In 2003, the LCS MM Program Office (PMS 420) was established as part of the Program Executive Officer (PEO) for Littoral and Mine Warfare (LMW). PMS 420’s responsibilities originally involved end-to-end LCS MP development and lifecycle support. In 2011, PEO LCS was established to align several program offices into one consolidated PEO focused entirely on delivering the LCS Program. This included the establishment of PMS 505 to assume the responsibility for Fleet Introduction and Sustainment of Mission Modules and the seaframes. PMS 420 continues to rely heavily on Participating Acquisition Resource Managers (PARMs) to deliver mission systems to meet LCS MP specific requirements and is effectively the lead SoS integrator of this complex system of manned and unmanned systems. Examples of PARM systems include manned helicopters, sensor systems, weapons systems, and unmanned aerial, surface, and sub-surface vehicles. In some cases where commercial or government technology solutions cannot be identified, PMS 420 works with the naval science and technology (S&T) community to conduct accelerated development of non-POR technologies.
The LCS MMs are being developed and delivered using a rapid fielding approach where initial increments are fielded as the PARM technology matures and is integrated into useful elements of capability. Key to enabling this incremental development has been the definition and configuration management of standard interfaces between the seaframe and the MPs, as defined in the Interface Control Document (ICD) for the LCS. This approach allows the seaframes and MMs to mature independent of each other’s development and facilitates reconfiguration of the LCS seaframe and incremental upgrade of the modular MPs over their life cycle. This independence of interfaces, which are managed by the common ICD, allows both programs to develop and test against their unique KPPs, key system attributes (KSAs), and milestones. This effectively enables the MP incremental fielding approach. Useful increments of capability are fielded, which build upon the initial capabilities, ultimately satisfying the full KPP requirements with the final increment. The incremental approach to delivering fleet capability also allows continued capability insertion throughout the life of the program. Capability Production Documents (CPDs) are developed to describe the capabilities required to meet the production baseline for a specific increment and also provide a testable requirement for the test community.

The LCS MM acquisition strategy leverages an organic Navy workforce, which provided flexibility to accommodate requirements refinement as the initial MP capability matured. This effectively reduced the Navy’s exposure to cost and contract risk with industry throughout the early prototyping phases. The various lead SoSEs are aligned to their respective chartered mission areas as shown in Figure 1 (i.e., Naval Surface Naval Surface Warfare Center [NSWC] Panama City Division is the LSoSE for the MCM MP, NSWC Dahlgren Division is the LSoSE for the SUW MP, Naval Undersea Warfare Center [NUWC] Newport Division is the LSoSE for the ASW MP, while SPAWAR Systems Center [SSC] Pacific provides overarching SoSE support for cross-package requirements and architecture activities).

Figure 1. LCS MM Lead SoS Systems Engineers by Mission Area

In a departure from the traditional prime integrator role, industry supports the LCS MM program as the Mission Package Integrator (MPI) to perform production, packaging, and assembly (PP&A) functions. The MPI role is limited and focuses on productionizing systems and equipment once a useful capability has been demonstrated. The various Naval Warfare and Systems Centers continue to assume primary responsibility of developing and integrating the initial MP increments through initial testing, fieldings, and early deployments.
as the LSoSE. The LCS MM acquisition workforce has also been supplemented significantly with functional area subject matter expertise as available from the various Navy Warfare and Systems Centers to provide cross MP support in areas which include, but are not limited to the following: System Safety, Environmental Safety and Occupational Health, Information Assurance/Cybersecurity, Configuration Management, Human Systems Integration, Corrosion Control, and so forth.

**LCS Mission Package and Open Architecture**

The LCS MMs program was founded on the principles of incremental acquisition, open system architecture, and an open business model. As previously discussed, incremental acquisition promotes fielding of needed capabilities as they become available. Standard interfaces and an open business model enable more rapid integration of PARM and other systems as they mature. The LCS MM reference architecture is shown in Figure 2. It is a high-level, system-agnostic block diagram that articulates functional as well as business dimensions. It is not intended to be the primary design tool, but is used by the lead SoSE and Program Manager (PM) to easily communicate SoS boundaries and interfaces with internal and external stakeholders, the S&T community, and other stakeholders. Additionally, it is used as the basis to establish boundaries for configuration management and certification events.

“Focused” warfare MPs are designed to be installed on an LCS one mission at a time with the LCS Combat System providing the MPs with access to the seaframe hosted information technology (IT) infrastructure, sensors, weapons, countermeasures, and communications reachback through Exterior Communications (EXCOMMS). Functionality provided by the MP Combat System includes, but is not limited to computing, communications, mission planning, execution, and post-mission analysis. MP manned and unmanned vehicles may be augmented with specific weapons, and/or sensors, and/or countermeasures to execute a focused mission. The level of detail of the reference architecture does not allow for the arrows to depict the specific data path, but implies connectivity at the end points.
Requirements, Architecture, and Interface Models

Orchestrating the development of a complex SoS requires clear communication and organization of requirements in ways that clearly derive and allocate requirements as well as define boundaries between where responsibility lies and where negotiation may be required between program offices, Systems Commands, or other stakeholders. The LCS MM specific example is further challenged by the need to manage multiple increments at varying levels of maturity and cost targets which drove concerted efforts to identify common services, hardware, and systems.

This requires a robust and disciplined SoSE activity to first decompose the high-level requirements communicated in the CDD and CPDs, and to then perform the detailed analysis required to support the technical and cost trades needed to ensure that the architecture, requirements, and interfaces are developed and maintained at a sufficient level of detail to ensure the development of a capable and cost effective SoS capability. A cross-MP Systems Engineering Integrated Product Team (SEIPT) was established at the working level responsible for developing requirements, architectures, and the associated SE tools and processes required to support the development of complete, verifiable, and cost effective functional and allocated baselines for the MPs which dictate the development of the LCS MM SoS. As various increments proceed through the development, PARMs and other external stakeholders are engaged as required to ensure that requirements and interfaces are understood and managed effectively. A series of Technical Interchange Meetings (TIMs) are typically held as a specific increment proceeds toward a formal Systems Engineering Technical Review (SETR; i.e., Preliminary Design Review [PDR]).
Critical Design Review [CDR], etc.). The LCS MM requirements framework was developed by the SEIPT and provides a hierarchical framework used to derive and allocate top-level requirements, prescribed by the CDD, down to the Mission Module and Mission System levels.

![LCS MMs Requirements Framework](image)

**Figure 3. LCS MMs Requirements Framework**

Level 1 requirements are defined as program level requirements provided by external sources. Level 1a requirements are developed by the user through the Joint Capabilities Integration Development System (JCIDS) process. For example, the LCS Flight 0+ CDD provides the required performance attributes for the initial MP Increments, and the Incremental CPDs provide the specific performance attributes in support of the incremental production and fielding of the MPs. Level 1b requirements define common external interfaces which are typically outside the span of PMS 420 or the LSoSE’s control. Level 2 requirements are defined as the common requirements developed by the LSoSE. These “common” or Level 2 requirements are typically driven by open architecture, cost savings, or other acquisition-type objectives. Level 2a captures sets of like functional requirements which are common to more than one Mission Package (i.e., Safety, Information Assurance/ Cybersecurity, Environmental, Logistics, etc.). Level 2b defines Systems and sub-systems which are common to more than one MP (i.e., computing infrastructure, communications, sustainment, etc.). The MP requirements are defined at Level 3 and typically include the level of detail to understand the incremental nature of the MP under development and/or
Mission Module requirements and intra-MM interfaces are typically defined at Level 4. Capabilities common to more than one MM within a specific MP, such as the MP Application Software (MPAS), are implemented at Level 4. Mission System requirements and interfaces are defined at Level 5. Interfaces which the LSoSE must retain control over to ensure capability or modularity objectives are documented in PMS 420 controlled Interface Requirement Specifications (IRSS) or ICDs.

The respective warfare focused MP technical teams manage the interfaces within an MM or MP by defining the details of the interfaces in the requirements and architecture products. Internal MP interfaces are documented within the standard section of the System/Sub-System Specification (SSS) and flowed down to Sub-system Specification (SS) and other documents. To control parallel development among the various MPs or external stakeholders, PMS 420 defines and controls common services and interfaces via standalone an IRS or an ICD/IDD depending on the level of control required. An IRS describes the functional requirements (e.g., data communication requirements) and an IDD or ICD describes the design details of the interface.

Program offices responsible for constituent system development typically contract to a prime developer and find it sufficient to decompose the CDD into a System/Sub-System Specification (SSS) or an A-Spec, which is then put on contract. Due to the complex nature of the LCS MM SoS and the number of organizations involved in development of MMs, mission systems, sub-systems, and so forth, PMS 420 has found it necessary to manage its specification tree down to the MM level (or Level 4) with a strong understanding of how those MM requirements are allocated to individual Mission Systems. PMS 420 has historically formalized such technical agreements with its PARMs or other stakeholders through System Project Directives (SPDs), which are essentially Memorandums of Agreement between government program offices supplemented with SSS, Sub-System Specification (SS), ICDs, and architecture products as appropriate. Architecture products are typically developed through the MM and mission system Levels 4 and 5 in accordance with version 2.0 of the Defense Architecture Framework (DODAF) framework (DoN, 2011). Interfaces which require high levels of integration and/or are anticipated to be high risk are further detailed using high fidelity interface models using Systems Modeling Language (SysML). Interface requirements further derived through the use of DODAF and SysML are fed back, or synchronized, into the appropriate requirements specification documents. The SPDs appended with SSS, SS, and IRSs at Level 4 or below to inform contract actions with PARM contract actions. As most of the mission systems which are constituent to the SoS have their own ACAT designation, requirements, and funding lines, it is not until this level of analysis is performed that we can begin to understand the gaps between the prescribed SoS and mission system requirements. This synchronized requirements and architecture development process is depicted in Figure 4.

The requirements derivation and supporting architecture development process requires collaboration across several organizations. PMS 420 utilizes a collaborative engineering environment such that its distributed user base can access the LCS MM DOORS modules from government and contractor facilities. In 2014, this capability was expanded to include access to the LCS MM System Architect encyclopedias. These DOORS and system architect instantiations are used mostly by the PMS 420 core SE IPT. In 2013, the Navy Systems Engineering Resource Center (NSERC) hosted Systems Engineering Integrated Data Environment (SE IDE) was redesigned to ensure internal and external users timely access to requirements, architectures, and technical data packages in support of formal reviews (i.e., SETRs, certification events, or other events as required).
Assessing Technical Performance Measures for a Complex SoS

Technical Performance Measures (TPMs) are a time-based metric used to compare actual performance to projected performance as a system advances through various developmental and test events. The specific metrics are typically implemented at the system or sub-system level. The TPMs which have the greatest impact on SoS mission performance are specified in the SPD and supporting requirements documents. The PARMs are responsible to report any scheduled test results back to the LSoSE on a quarterly basis, or sooner via the integrated risk management, or other technical management process. The LSoSE is then challenged with understanding the impact of any performance shortfalls and then assessing the impact to SoS mission performance and presenting the impacts to leadership.

The LCS MM program developed a framework to identify and track the TPMs which are critical for LCS mission success. Therefore, thoughtful selection, review, and tracking of TPMs are required to allow technical managers to make informed decisions during system design and to identify the need for corrective actions when deviations from planned technical progress occur. In cases of complex SoS with multiple variables, dependencies and/or mission execution paths, multiple TPMs inform a modeling and simulation (M&S) tool to project probable performance of the MSs and/or MMs associated with an MP. In some cases, the TPMs that PARMs are tracking for other missions are not critical to PMS 420 and therefore are not tracked by the LCS MM LSoSE.

The allocation of KPPs to TPMs for a notional ASW MP is depicted in Figure 5. High fidelity, non-linear, complex M&S tools are usually required to predict the performance of the integrated MP based on the TPMs’ inputs of the constituent mission systems. For example,
the Naval Mine Warfare Simulation (NMWS) tool is an object-oriented, event-driven simulation tool traditionally used by the Navy’s Mine Warfare community. This SoS M&S tool is capable of reflecting the complexity of an SoS because it can accommodate multiple inputs, such as mission thread information, environmental predictions, mission system performance parameters, high-level reliability parameters, and critical task analysis inputs. PMS 420 has found the use of a Kill Chain based modeling methodology to provide benefit to program management and risk identification, and we will now discuss that process.

Applying the Kill Chain Methodology to a Complex SoS

Given a notional SoS with 12 constituent systems, each with six to eight TPMs, over 100 TPMs can result. The LSoSE cannot always expect to present status on each TPM at TIMs, SETRs, certification events, or other decision meetings. To assist in managing this challenge, the Kill Chain methodology of a mission-based approach to assessing and visualizing how an SoS is expected to perform in the execution of a specific mission thread is used. It is intended to provide insight into any risks or issues associated for the individual systems, sub-systems, or interfaces in properly executing its allocated functionality. Several communities have developed specific implementations for various purposes. Perhaps the most pervasive are the operational test community’s Warfare Capability Baseline (WCB) process and larger Navy Integration and Interoperability (I&I) implementations (Clawson et al., 2015). The WCB approach in particular was developed with a focus on test data in preparation for an operational test event. LCS MM LSoSEs have expanded upon and tailored a Kill Chain implementation to compensate for the incremental, technical, and organizational complexities that can complicate the assessment and presentation for a complex SoS.

As previously described, the LCS MM SoS is made up of various manned and unmanned vehicles, sensor systems, communication systems, weapons systems, and systems provided by the core LCS seaframe. As MPs are fielded incrementally, mission systems and/or capability upgrades are added to the MP baselines. Further, multiple sub-
mission threads, or “passes,” with various permutations of these vehicles, sensors, and systems, can be required to accomplish missions with extended timelines, such as those required to clear a minefield. An example Kill Chain for a notional ASW MP is shown in Figure 6. The overarching Kill Chain assessment process is depicted in Figure 7. The Kill Chain is presented as a matrix of systems and sub-systems, laid against mission phases (i.e., Prepare/Configure, Search/Detect/Classify/Localize, Identify, Neutralize, Assess, Post Mission Analysis). Each system, sub-system, and integration point is assessed for its individual contribution associated and risk of meeting the mission specific requirements set using the LCS MM standard risk assessment criteria. In many cases, multiple TPMs are tracked and used to quantitatively assess the respective system or sub-system against its allocated contributions. Any system which misses a major TPM milestone cues a re-evaluation of the Kill Chain for impact. If the performance degradation does not force a change to the risk rating, then the Kill Chain is not updated. The assessment method is also denoted to provide insight into the level of confidence in the assessment (i.e., S = SME estimate, M = Modeled result, T = Verified in test, F = Fielded System). A path key with varying thickness is used to assist the PM or LSoSE to understand where the specific system falls in the multi-pass sequence of events. Notes can be overlaid on the Kill Chain to reference the PM or LSoSE to specific risks or issues. In order to not lose the fidelity, Kill Chains are typically created for major test configurations, deployed configurations, or Increments.

Figure 6. Example Kill Chain for a Notional ASW MP
Lessons Learned and Opportunities for Future Work

This paper presented the systems engineering processes and level of technical rigor required to manage the development and integration of a complex SoS. A specific Kill Chain framework was presented, which is currently used to assess and visualize the impact of TPMs on a complex systems of manned and unmanned systems mission threads. As currently constructed, the Kill Chains provide the PM with comprehensive insight into the technical status and major risks and issues for a complex SoS across specific mission threads. Due to the various configurations that the LCS MM PMs and LSoSEs are responsible to track, the LCS MM program office has invested significant resources into coordinating several of the basic systems engineering methodologies and tools to support the Kill Chain process, namely requirements, architectures, and risk management.

- Requirements were developed to a sufficient level of detail to articulate the specific contribution of the system, sub-system, and interface required at each phase in the mission thread.
- Architectures were developed to the sufficient level of detail to ensure that the weapons critical path covers the sequence of events critical for mission success.
- Constituent system requirements must be aligned to LCS MM requirements. In the cases when constituent systems have their own requirements sets
(CDDs, CPDs, or other), the LSoSE must work with both resource sponsors to identify and work to close significant requirements and/or interface misalignments.

- Constituent system program managers must also be responsible to, and have a mechanism to report, TPMs to the LSoSE routinely.
- Risk and issues are flagged at least to the level of detail (i.e., system, sub-system, interface) presented in the Kill Chain.
- The team responsible to develop MM requirements is distributed across at least four Naval Warfare and Systems centers. PMS 420 has developed a common access card (CAC)-enabled collaborative engineering environment where a distributed user base can access the LCS MM DOORS modules from government and contractor facilities. Change management and control processes were also developed. In 2014, this capability was expanded to include access to the LCS MM System Architect encyclopedias.

To date, the Kill Chain framework has been extremely effective as LCS MM prepares for several operational test events. Performance issues are captured, assessed, and presented in the context of their impact to the overall mission success. In order to expand this approach to provide the PM with a more predictive tool that can be used for technology insertion planning and to better support the developmental phases of the next increments, some of the other technical management and systems engineering methodologies and tools must be appropriately configured. Integrating a robust modeling and simulation capability would better support analysis of alternative and technology insertion trade-off–type analyses. Integrating cost, schedule, and the -ilities, such as reliability, sustainability, and other models, would better support understanding the capability from a performance and total ownership cost (TOC) perspective.

References


