NPS-AM-26-227

ACQUISITION RESEARCH PROGRAM SPONSORED REPORT SERIES

Best Practice Alignment and Program Outcomes in DoD Middle Tier Acquisition of Acquisition: A Cross-Program Analysis

June 2026

CPT Chandra Meadows, USA

MAJ Matthew B. Teeple, USA

Thesis Advisors: Dr. Robert F. Mortlock, Professor
Jeffrey R. Dunlap, Lecturer

Department of Acquisition, Finance and Manpower

Naval Postgraduate School

Prepared for the Naval Postgraduate School, Monterey, CA 93943

The research presented in this report was supported by the Acquisition Research Program of the Department of Acquisition, Finance and Manpower at the Naval Postgraduate School.

ABSTRACT

This research examines the extent to which Department of Defense (DoD) Middle Tier of Acquisition (MTA) rapid prototyping programs align with selected acquisition best-practice constructs. The study employed structured content analysis using selected program artifacts from the Defense Acquisition Visibility Environment (DAVE). The program artifacts were quantitatively analyzed using the GenAI.mil large language model (LLM) to ordinally score and provide document references for the scoring rationale. The rubric-constrained scoring model was applied to evaluate six best-practice constructs: Modular Open Systems Approach (MOSA), iterative development, modeling and simulation (M&S), delegation of authority, business case quality, and technological readiness level (TRL) at program entry. The outputs were cataloged and analyzed using descriptive comparisons and exploratory regression. Descriptive comparisons evaluated differences between successful and unsuccessful programs, Service-level differences, and acquisition document-type analysis in relation to the best-practice constructs. Exploratory regression was used to investigate relationships among the best-practice construct variables, including selected multivariate models. This research contributes a structured framework for evaluating MTA program structures and provides recommendations to improve acquisition visibility, assessment, and decision-making within rapid prototyping efforts.

THIS PAGE INTENTIONALLY LEFT BLANK

ABOUT THE AUTHORS

CPT Chandra Meadows is an Acquisition Officer in the United States Army. She graduated from East Carolina University in 2010 where she received a Bachelor of Science in Business Management. She commissioned through Officer Candidate School in Fort Benning, Georgia. After graduating from the Naval Postgraduate School she will be reporting to Fort Campbell, KY to be a Contracting Officer.

MAJ Matthew Teeple is an Acquisition Officer in the United States Army. He graduated from Florida State University in 2015 and received a Master of Science in Urban and Regional Planning. He enlisted in the Army National Guard in 2008 and received a commission via Army ROTC in 2015. After graduating from the Naval Postgraduate School, he will report to Fort Carson, CO, to serve as a Contracting Officer.

THIS PAGE INTENTIONALLY LEFT BLANK

NPS-AM-26-227

ACQUISITION RESEARCH PROGRAM
SPONSORED REPORT SERIES

Best Practice Alignment and Program Outcomes in DoD Middle Tier Acquisition of Acquisition: A Cross-Program Analysis

June 2026

CPT Chandra Meadows, USA

MAJ Matthew B. Teeple, USA

Thesis Advisors: Dr. Robert F. Mortlock, Professor
Jeffrey R. Dunlap, Lecturer

Department of Acquisition, Finance and Manpower

Naval Postgraduate School

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AAF	Adaptive Acquisition Framework
ADM	Acquisition Decision Memorandum
AIR	Acquisition Information Repository
AoA	Analysis of Alternatives
BES	Budget Estimate Submission
CAE	Component Acquisition Executive
CBRN	Chemical, Biological, Radiological, Nuclear
CJCS	Chairman of the Joint Chiefs of Staff
CRC	Concept Requirement Capability
CTE	Critical Technology Elements
CV	Combat Evaluation
DA	Decision Authority
DAE	Defense Acquisition Executive
DAS	Defense Acquisition System
DAU	Defense Acquisition University
DAVE	Defense Acquisition Visibility Environment
DBS	Defense Business Systems
DFARS	Defense Federal Acquisition Regulation Supplement
DoD	Department of Defense
DoDAF	Department of Defense Architecture Framework
DoDD	Department of Defense Directive
DoDI	Department of Defense Instruction
DoDIG	Department of Defense Office of Inspector General
DOTMLPFP	Doctrine, Organization, Training, Materiel, Leadership, Personnel, Facilities, and Policy
DPG	Defense Planning Guidance
EM	Electromagnetic
EMD	Engineering and Manufacturing Development
FAR	Federal Acquisition Regulation
FDD	Foundation for the Defense of Democracies
FG	Fiscal Guidance

FY	Fiscal Year
GAO	Government Accountability Office
GenAI	Generative Artificial Intelligence
INCOSE	International Council on Systems Engineering
IT	Information Technology
IT-AAC	Information Technology Acquisition Advisory Council
ITRA	Integrated Technical Risk Assessment
JFRP	Joint Force Requirements Process
JOP	Joint Operational Problem
LCSP	Life Cycle Sustainment Plan
LLM	Large Language Model
M&S	Modeling and Simulation
MCA	Major Capability Acquisition
MDA	Milestone Decision Authority
MOSA	Modular Open Systems Approach
MTA	Middle Tier of Acquisition
MTARF	Middle Tier of Acquisition Rapid Fielding
MTARP	Middle Tier of Acquisition Rapid Prototyping
NDAA	National Defense Authorization Act
NDIA	National Defense Industrial Association
NDS	National Defense Strategy
NMCARS	Navy Marine Corps Acquisition Regulation Supplement
NMS	National Military Strategy
OMB	Office of Management and Budget
OUSD(A&S)	Office of the Under Secretary of Defense for Acquisition and Sustainment
OUSD(R&E)	Office of the Under Secretary of Defense for Research and Engineering
PDR	Preliminary Design Review
PM	Project Managers
POM	Program Objective Memorandum
PPBE	Planning, Programming, Budgeting, and Execution
RDT&E	Research, Development, Test, and Evaluation

RF	Rapid Fielding
RP	Rapid Prototyping
RWG	Rapid Working Group
SAMP	Simplified Acquisition Management Plan
SECDEF	Secretary of Defense
SORRD	Special Operations Rapid Requirements Document
TRA	Technology Readiness Assessment
TRL	Technology Readiness Level
UCA	Urgent Capability Acquisition
USSOCOM	United States Special Operations Command
WAS	Warfighting Acquisition System
WBS	Work Breakdown Structures

THIS PAGE INTENTIONALLY LEFT BLANK

ARTIFICIAL INTELLIGENCE DISCLOSURE

The authors used generative artificial intelligence (AI) tools in a limited and controlled manner to support aspects of this research, with prior coordination and approval from the thesis advisor. GenMIL AI was used to assist in the process of artifact scoring process across collected program documentation. The scoring variables were defined by the authors, while generative AI tools were used to help structure the scoring criteria, evaluation framework, and associated prompts to promote consistency in application. The authors maintained full control over the scoring process, interpretation of results, and all conclusions presented in this thesis.

OpenAI's ChatGPT (GPT 5.3) was used to assist in refining and structuring prompts utilized within GenMIL AI and to support the authors' understanding of statistical outputs during the regression analysis phase. This support was limited to explanatory and educational purposes to enhance the authors' interpretation of results. ChatGPT did not generate research findings or make analytical decisions.

All outputs generated through AI-supported processes were critically reviewed and validated by the authors to ensure alignment with the defined research design, source materials, and analytical intent. The authors maintained full responsibility for the methodology, data integrity, analysis, and conclusions presented in this thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The Department of Defense (DoD) has emphasized speed as a critical requirement for modern acquisition, recognizing that traditional major capability acquisition processes are insufficient to meet the demands of rapidly evolving threats. Recent guidance and acquisition reform efforts highlight that delivering capability at pace is no longer optional but essential to maintaining operational advantage. As articulated by the Secretary of War, Pete Hegseth, “our objective is simple: transform the entire acquisition system to operate on a wartime footing,” underscoring the urgency to align acquisition practices with the tempo of modern warfare (Phillips, 2025, para. 24).

While this emphasis on speed reflects current strategic priorities, it is not a new challenge within defense acquisition. The DoD established the Middle Tier of Acquisition (MTA) pathway, designed to accelerate capability development and deliver operationally relevant systems within compressed timelines. As MTA adoption has expanded, it has become a mechanism for enabling rapid innovation and fielding across a wide range of programs. However, outcomes across MTA rapid prototyping efforts have demonstrated significant variability.

A. PROBLEM STATEMENT

Existing oversight and research, including assessments conducted by the Government Accountability Office (GAO) and industry, identify a set of best practices associated with successful program outcomes. These include the use of Modular Open Systems Approach (MOSA), achievement of appropriate Technology Readiness Level (TRL) at program entry, development of a sound business case, effective use of modeling and simulation, iterative development, and delegation of authority, as illustrated in Figure 1.

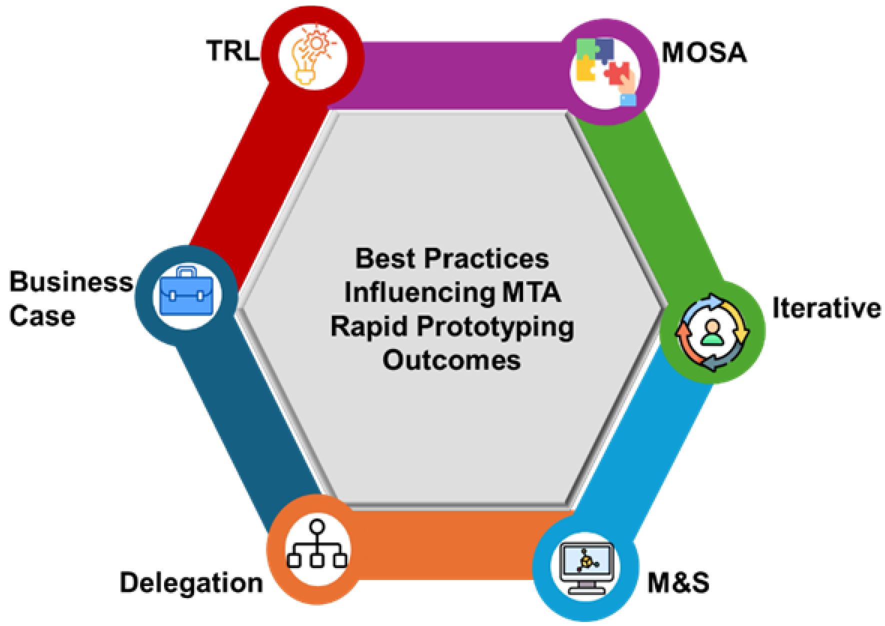

Figure 1. Middle Tier of Acquisition Best Practices Framework

Despite the identification of these best practices, existing research remains largely program-specific and does not provide a systematic, cross-program assessment of their implementation within MTA efforts. As a result, there is limited empirical understanding of whether alignment with acquisition best practices constructs influences success within MTA rapid prototyping programs. The reliance on program documentation often reflects stated intent rather than actual execution, making it difficult to determine whether observed outcomes are driven by adherence to best practices or by other contextual factors.

This gap is particularly significant given the increasing reliance on MTA as a primary pathway to accomplish speed. Without a structured, data-driven assessment of how best practices are applied across programs, acquisition professionals lack the ability to identify patterns, anticipate risk, or make informed decisions regarding pathway selection and program execution. If MTA rapid prototyping efforts are implemented without a clear understanding of the conditions under which they are most effective, the DoD risks misallocating resources, extending timelines, and delaying the delivery of critical capabilities to the warfighter.

Therefore, this research addresses the existing gap through an assessment of DoD MTA rapid prototyping efforts against acquisition best practices and an analysis of how adherence to those practices corresponds with program outcomes.

B. RESEARCH QUESTIONS

This study examines the factors that influence the success and failure of MTA rapid prototyping efforts across the DoD. While the MTA pathway was designed to accelerate capability development, variation in program outcomes suggests that successful execution depends on how programs implement best practices.

1. Primary Research Question

- To what extent do DoD MTA rapid prototyping programs demonstrate documented alignment with selected acquisition best-practice constructs, and how does that alignment vary across program status categories, Services, and best-practice domains?

2. Secondary Research Questions

- How consistently are selected acquisition best practice constructs documented across MTA rapid prototyping program artifacts in the AIR/DAVE dataset?
- How does documented best-practice alignment differ across AIR/DAVE program status categories, including programs identified as successful, transitioned, suspended, terminated, disposed, or otherwise unsuccessful?
- How does documented best-practice alignment vary across military Services within MTA rapid prototyping programs?
- What exploratory relationships exist among documented best-practice constructs within MTA rapid prototyping program documentation?

C. RESEARCH OBJECTIVES

This study evaluates the extent to which MTA rapid prototyping programs align with acquisition best practice constructs and how this alignment relates to program outcomes across the DoD. Using data from the Defense Acquisition Visibility Environment (DAVE), the research assesses the presence and consistency of these practices across a set of MTA rapid prototyping programs.

The study further compares the implementation of best practices between successful and unsuccessful programs to identify differences in execution patterns and determine which practices demonstrate the strongest relationships with successful outcomes.

The findings aim to provide actionable insights that improve the alignment between acquisition practices and program outcomes, thereby supporting more effective decision-making within the MTA rapid prototyping framework.

D. METHODOLOGY

This study employs a structured qualitative content analysis approach to evaluate the extent to which MTA rapid prototyping programs align with acquisition best practice constructs and how that alignment relates to program outcomes. The analysis draws on program documentation contained within the Acquisition Information Repository (AIR) of DAVE, which provides a centralized source of acquisition artifacts across DoD programs.

The best practice framework was developed using GAO guidance spanning from the implementation of the MTA pathway to the present, with additional validation through defense acquisition literature, policy sources, and industry best-practice concepts, identifying six key practices. These practices form the basis for evaluating program documentation and assessing the extent to which they are reflected across MTA rapid prototyping efforts.

To enable analysis across a large and variable dataset, this study incorporates large language model (LLM)-assisted analysis using the GenAI.mil Gemini model to support the systematic review of program artifacts. Structured prompts were designed to constrain the model to extract only explicit textual evidence, minimizing inference and improving consistency in evaluation. Program artifacts were individually assessed and scored based on the presence and strength of best practice indicators, enabling standardized comparison across programs.

After scoring all individual program level artifacts, the resulting dataset was aggregated to enable analysis across multiple analytical dimensions. The document level ordinal scores were averaged to produce program level measurements for each of the best practice constructs, resulting in a structured dataset that captures the presence and strength of each best practice construct. These aggregated measures enable comparison across document types, military services, program outcomes, and between best practice constructs.

This analysis takes place in two main areas. First, descriptive analysis is conducted to examine best practice implementation patterns across acquisition document types and between aggregated successful and unsuccessful programs. It further includes a comparative assessment of Service-level variation within both successful and unsuccessful programs. These comparisons provide insight into how acquisition practices are distributed across programs and Services and where variation in documentation occurs.

The second is a linear regression analysis used to explore relationships among the best practice constructs. These regression models assess the extent to which variation in one construct can be associated with variation in another, providing insight into how these acquisition practices interact with the selected MTA rapid prototyping efforts. This regression analysis is exploratory in nature and is used to evaluate the strength, direction, and consistency of relationships across the best practice constructs, and not to establish causation.

This dual-stage analytical approach attempts to integrate qualitative content analysis with quantitative methods, enabling an assessment of MTA rapid prototyping program efforts and their relationship to the best practice framework.

E. SCOPE AND LIMITATIONS

This study focuses on MTA rapid prototyping programs across the DoD, utilizing data available within DAVE. The analysis is limited to programs that provide sufficient documentation and emphasizes the evaluation of program alignment with acquisition best practice constructs. The scope of the research is restricted to unclassified and open-source documentation to ensure compliance with data access and distribution requirements.

Several limitations influence the findings of this study. First, the availability and consistency of documentation within DAVE are uneven across programs. Program records are often incomplete, sporadic, or concentrated in early phases, with limited documentation capturing full program execution or closeout assessments. This inconsistency constrains the ability to conduct full lifecycle analysis and may impact the completeness of cross-program comparisons.

Second, the reliance on documented evidence may reflect program intent rather than actual execution, limiting the ability to definitively assess how best practices were implemented in practice.

Third, technical constraints associated with data processing limit the ability to employ scaled automated data processing scripts, such as Python-based workflows. As a result, the study relied on a combination of manual review and LLM-assisted analysis, which may limit scalability and introduce variability in interpretation.

Finally, the dataset contains a limited number of clearly identifiable unsuccessful or failed programs, which constrain the ability to conduct a fully balanced comparative analysis. This limitation may affect the strength of conclusions regarding the relationship between best practice implementation and program failure.

F. SUMMARY

Chapter I established the foundation for examining the effectiveness of MTA rapid prototyping programs within the context of the DoD's increasing emphasis on acquisition speed. It identified a critical gap in understanding how acquisition best practices are applied across programs and how that alignment relates to program outcomes.

In response, this study defined research questions and objectives centered on evaluating program alignment with acquisition best practice constructs using data from DAVE, while acknowledging key methodological and data limitations. By framing the analysis around cross-program evaluation and evidence-based assessment, this research seeks to provide insights that improve acquisition decision-making and program execution.

Building on this foundation, Chapter II will examine the historical development of the MTA pathway and the broader acquisition environment in which it operates.

II. BACKGROUND

This capstone research provides the reader with sufficient information to understand in detail how MTA fits within the broader Warfighting Acquisition System (WAS) framework, how programmatic efforts are undertaken within this framework, and the different military departmental approaches to operationalizing MTA within their respective services.

A. THE WARFIGHTING ACQUISITION SYSTEM

The term *Warfighting Acquisition System* has dual meanings in the U.S. defense context; Figure 2 illustrates this duality by depicting both systems.

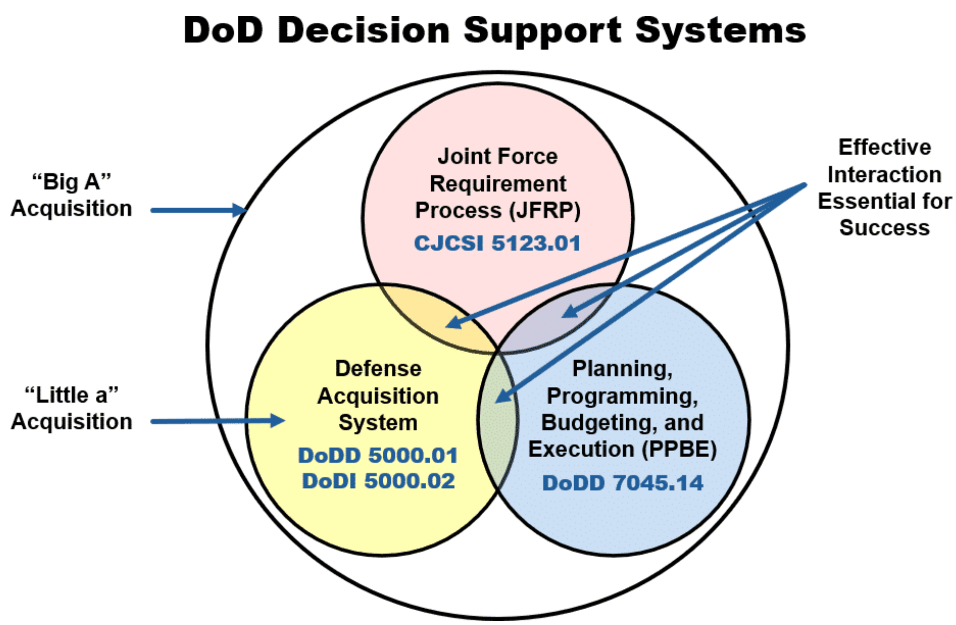

Figure 2. The Warfighting Acquisition System. Source: Defense Acquisition University (2026)

The first meaning of the Warfighting Acquisition System is referred to as "big A" Acquisition, which is a representative whole-of-effort construct that explicitly includes the Joint Forces Requirements Process (JFRP); planning, programming, budgeting, and execution (PPBE), and Systems Acquisition Management, also commonly referred to as the Adaptive Acquisition Framework, as depicted in Figure 2. The JFRP is responsible for articulating required capabilities by translating joint warfighter needs into a prioritized list of Joint Operational Problems (JOPs) and their associated Concept Required Capabilities (CRCs) (Chairman of Joint Chiefs of Staff, 2026, p. A-7). JFRP uses strategic planning documents like the National Defense Strategy (NDS) and the National Military Strategy (NMS), along with input from the Chairman of the Joint Chiefs of Staff (CJCS), Combatant Commanders (CCDRs), experts, and adjacent stakeholders, to develop JOPs, along with an analysis of how to translate them into CRCs (CJCS, 2026, p. A-3, A-4). CRCs are organized using the Doctrine, Organization, Training, Materiel, Leadership, Personnel, Facilities, and Policy (DOTmLPFP) framework, which examines a range of non-materiel solutions in addition to a commercial materiel solution vector (CJCS, 2026, p. A-7).

If a materiel solution requiring development or integration is determined to be the most practical option and validated by the requirements process as a materiel need, the new requirement is assessed for funding through the PPBE process (McGarry, 2022, p. 5-6). The PPBE process is a calendar-driven sequence of events executed to secure funding for an acquisition program solution to a materiel requirement identified through the JFRP process (McGarry, 2022, p. 2, 10). The planning phase is broadly defined as the process of contextualizing the strategic importance of the capabilities requirements identified by JFRP into objectives for programming at the service staff level, along with input from the Defense Planning Guidance (DPG) issued by the Secretary of Defense (SECDEF) (Candreva, 2023, p. 227). The programming phase is where resources are allocated across existing and new programs. These resources are balanced against fiscal guidance (FG), which is issued as memoranda by the White House Office of Management and Budget (OMB) (Candreva, 2023, p. 228). While the FG is not the president’s budget, it serves as a guidance document for service and defense planners who, in turn, submit their Budget Estimate Submission (BES) based on the figures provided in the FG and align these funds with the Program Memorandum Objective (POM) (McGarry, 2024). The POM is then used to execute a budget review, during which funding allocations are reviewed and ultimately sent to Congress for enactment (Candreva, 2023, p. 247). Execution is the final phase, in which authorized program elements execute their enacted budgets to pursue the objectives laid out in the POM, and where PPBE blends with the system acquisition management, often referred to as “little a” acquisition (Candreva, 2023, p. 247-248).

The Warfighting Acquisition System is the process by which materiel capabilities are built, bought, fielded, and sustained throughout their lifecycle. Its principles and tenets are laid out in Figure 3 and are operationalized through the Adaptive Acquisition Framework (AAF).

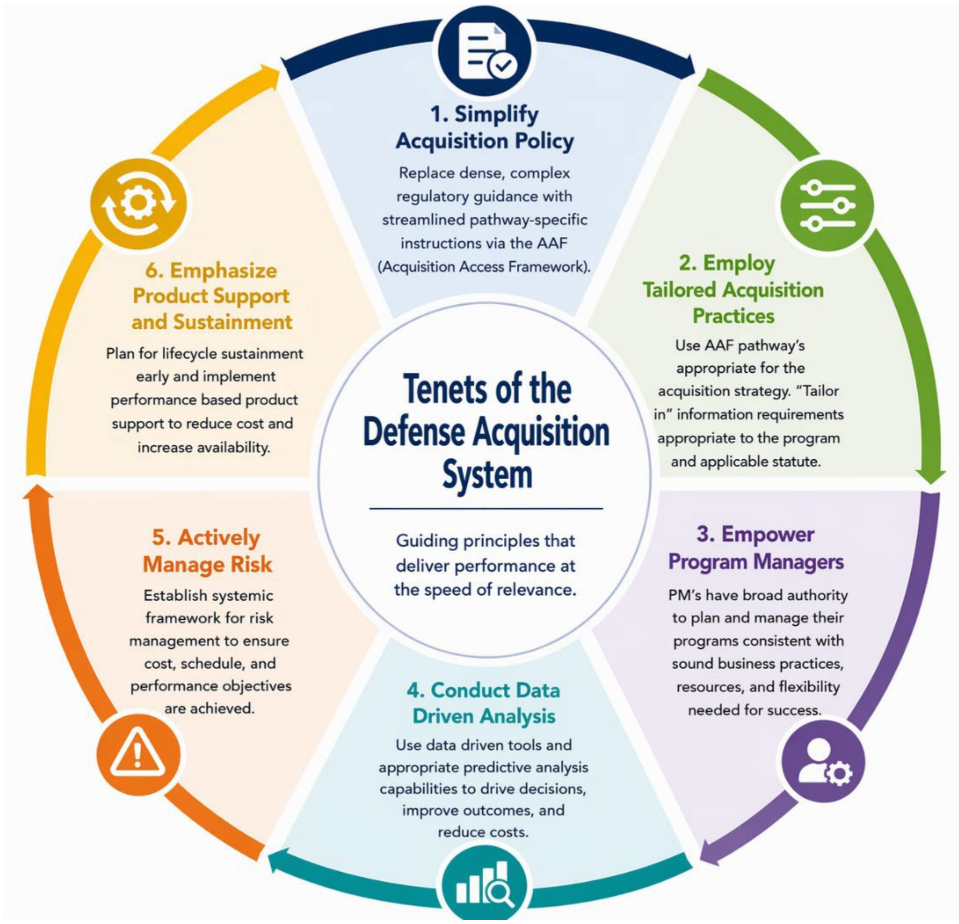

Figure 3. Tenets of the Defense Acquisition System. Adapted from Office of the Under Secretary of Defense for Acquisition and Sustainment (2020b); Office of the Under Secretary of Defense for Acquisition and Sustainment (2020a)

B. THE ADAPTIVE ACQUISITION FRAMEWORK

The AAF replaced the previously titled Operation of the Defense Acquisition System, which was a milestone-based acquisition governance model that could be modeled in several different ways based on programmatic characteristics (Office of Undersecretary of Defense for Acquisition, Technology, and Logistics [OUSD[AT&L]], 2017, pp. 11-17). Similar to the AAF that succeeded it, the Operation of the Defense Acquisition System, in its most refined and final 2017 form, encouraged the tailoring of a specific program based on considerations of product type and the speed at which the development effort was needed by the warfighter (OUSD[AT&L], 2017, p. 9). This framework also provided an ability to enter at advanced milestones (i.e., B or C), combine or adjust decision events, and align oversight intensity with a program's overall level of technological maturity, urgency of need, and operational context (OUSD[AT&L], 2017, pp. 3, 19). However, this earlier framework remained primarily organized around milestone alignment, sequential acquisition phases, and formal decision gates, governed by milestone decision authorities; it did not yet establish distinct acquisition pathways later formalized under the AAF (OUSD[AT&L], 2017, pp. 11–19; OUSD[A&S], 2020a, p. 12).

In January 2020, the Office of the Under Secretary of Defense for Acquisition and Sustainment (OUSD[A&S]) released a comprehensively updated acquisition framework that incorporated requirements from adopted legislation enacted under the National Defense Authorization Acts (NDAA). In particular, Section 804 from the fiscal year (FY) 2016 National Defense Authorization Act (NDAA), which mandated the creation of the "Middle Tier of Acquisition for Rapid Prototyping and Rapid Fielding," (National Defense Authorization Act for Fiscal Year 2016 [NDAA 2016], 2015) and Section 873 of the FY 2018 NDAA titled "Pilot Program to Use Agile and Iterative Development Methods to Tailor Major Software-Intensive Warfighting Systems and Defense Business Systems" (National Defense Authorization Act for Fiscal Year 2018, 2017). Section 804 would become the Middle Tier of Acquisition, and Section 837 would ultimately become the Software Acquisition Pathway within the AAF.

The AAF is intended to meet the needs of the WAS specifically as outlined in DoDD 5000.01 and its specific mandate to "Deliver Performance at the Speed of Relevance" (OUSD[A&S], 2020c, p. 4). It seeks to do so through six tailorable pathways that offer significantly greater flexibility than its predecessor, and precepts tailored to the acquisition type being initiated. Figure 4 lists these pathways as constructed in DoDI 5000.02.

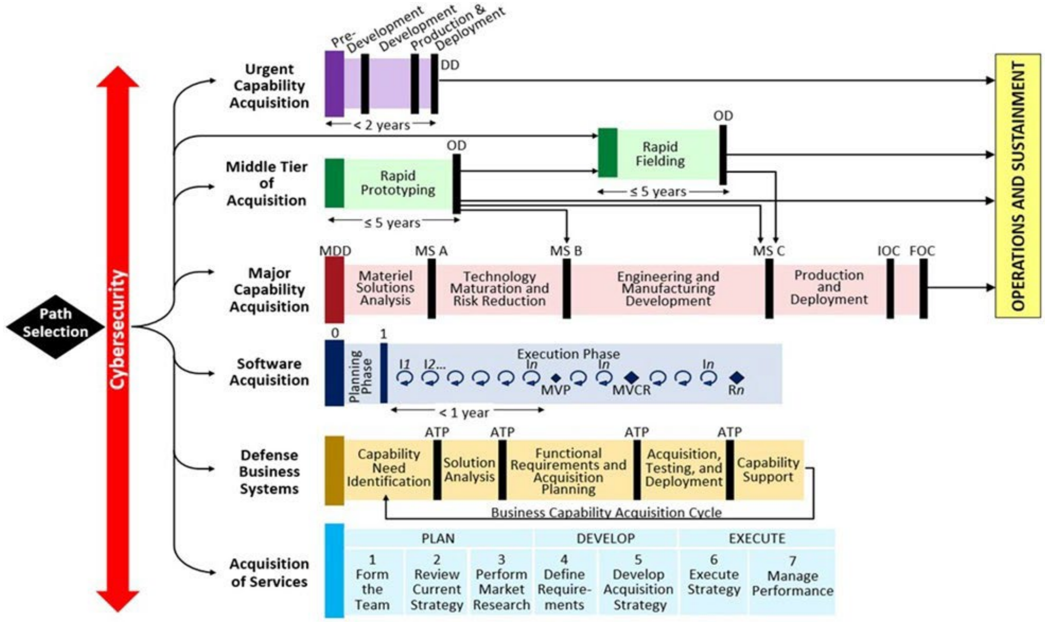

Figure 4. Adaptive Acquisition Framework. Source. Under Secretary of Defense for Acquisition and Sustainment (2020b), p. 10

In brief, the pathways adopted include Urgent Capability Acquisition (UCA), which provides a rapid pathway for addressing urgent or emergent operational needs, with an objective of delivering a suitable capability in under two years (OUSD[A&S], 2020a, p. 12). The MTA pathway is designed for capabilities mature enough to move quickly, either through rapid prototyping or rapid fielding (DAU, 2023). Under the rapid prototyping path, the objective is to produce fieldable prototypes that demonstrate novel and/or residual operational capabilities within five years (OUSD[A&S], 2020a, p. 13). Under the rapid fielding path, proven technologies requiring limited development and demonstrating sufficient manufacturing readiness to start production within six months, and fielding completed within five years (OUSD[A&S], 2020a, p. 13). Major Capability Acquisition (MCA) seeks to develop, acquire, and modernize capabilities of an enduring nature (OUSD[A&S], 2020a, p. 14). The Software Acquisition Pathway enables incremental delivery of software capabilities using modern development approaches, such as Agile, DevSecOps, and Lean Practices (OUSD[A&S], 2020a, p. 14). Defense Business Systems (DBS) Acquisition integrates services acquisition to acquire information technology solutions that have been successfully demonstrated in the commercial marketplace with minimal modifications (OUSD[A&S], 2020a, p. 15). Finally, the

Acquisition of Services Pathway is meant to guide the identification and acquisition of defense-related services (OUSD[A&S], 2020a, p. 16).

C. OBJECTIVES OF MIDDLE TIER ACQUSITION

The concept of time-to-market and its inverse relationship to product performance within a given marketplace is well explored in the business community (Cohen et al., 1996, pp. 173–174). A keen understanding of the trade-offs, associated risks, and opportunity windows is essential to the planning and execution of product delivery relevant to a competitive marketplace (Cohen et al., 1996, pp. 173–174, 176). A core tension in innovative marketplaces is that accelerating a given system's performance typically requires a longer development timeline, yet longer development times may cause a firm to miss a finite window of opportunity in the marketplace. Deliberate tradeoff management is a central component of competitive product delivery (Cohen et al., 1996, pp. 174, 178).

These fundamental marketplace frictions are what MTA hopes to solve within a defense acquisition context. MTA governance puts a firm ceiling on time-to-market, capping it at 5 years for both Rapid Fielding and Rapid Prototyping (OUSD[A&S], 2019, p. 4). This fixed development window operationalizes the "delivery of performance at the speed of relevance" mantra by subordinating the length of the development cycle to the relevance of the opportunity window (OUSD[A&S], 2019). MTA addresses and attempts to fill the gap for capabilities that possess adequate technological and manufacturing maturity to be rapidly prototyped or fielded within five years of program initiation, and as a mechanism to accelerate maturation or minimally develop before fielding (OUSD[A&S], 2019, p. 4).

MTA does have limitations and is generally not appropriate for major systems projects intended for interagency use, projects with significant international partner involvement, or projects primarily focused on technological development (OUSD[A&S], 2019, p. 5). These would typically be better suited for Major Capability Acquisition or, dependent on the technology in question, the Software Acquisition pathway (OUSD[A&S], 2020a).

D. THE INS AND OUTS OF MIDDLE TIER ACQUISITION

The MTA pathway provides a "middle tier" approach somewhere between the urgent and deliberate approaches to acquisition and system development (DAU, 2023, p. 9). It delivers capabilities to the end user in a timescale of greater than two years but fewer than five (DAU, 2023, p. 9).

Implementation of the Adaptive Acquisition Framework (AAF) opened the door for DoD program managers to leverage multiple acquisition pathways, both concurrently and sequentially, to deliver capabilities not otherwise obtainable through a single pathway (OUSD[A&S], 2020a, p. 10). Rather than being bound to the constraints of a program of record using the major capability acquisition (MCA) pathway, program offices may now undertake concurrent but discrete efforts using the same pathway, for example, two or more unique prototyping efforts to yield innovative and viable solutions using the MTA pathway (GAO, 2023a, p. 6). One of the unique aspects of the MTA pathway is its ability to serve as a springboard for such prototyping and fielding and then transition these initial programmatic efforts into the MCA pathway (DAU, 2023, p. 13).

This flexibility is especially relevant for the Middle Tier of Acquisition Rapid Prototyping (MTA RP) pathway, given its structural schedule constraints that limit Rapid Prototyping initiatives to 5 years (OUSD[A&S], 2019, p. 4). Follow-on efforts could benefit from transitioning to MCA or to Rapid Fielding, also under the broader scope of MTA (DAU, 2023, p. 13). Rapid prototyping can accelerate capability development and reduce technical risk before transitioning to another pathway (DAU, 2023, p. 13). Completing a transition anticipation, development, and coordination of information equirements to define the transition points from one pathway to another ultimately informs the exit criteria and documentation required (OUSD[A&S], 2020a, p. 11).

Concurrent development under the AAF, specifically between the Software Pathway and other, more hardware-centric pathways such as MCA or MTA, is also explicitly authorized (OUSD[A&S], 2020a, pp. 8-9). Such development potentially reduces risk in system architecture by enabling iterative software development through the Software Pathway, while the MCA or MTA hardware architecture is informed by the system definition characteristics it must meet to host and integrate the co-developed software into a shared interface platform (OUSD[A&S], 2020c, p. 8; INCOSE, 2015, pp. 71-74, 79–81). Additionally, the iterative nature of software means its update cycle is significantly more frequent than hardware updates, placing both in the same programmatic effort and timeline, which has proven problematic in the past (DeLuca et al., 2013, pp. xxi–xxiv).

E. MAJOR AND NON-MAJOR MIDDLE TIER ACQUISITIONS

MTAs are categorized as "major" and "non-major" MTAs (OUSD[A&S], 2019, p. 13). This distinction is delineated by the total expected expenditure of the MTA program (OUSD[A&S], 2019, p. 13). If an MTA program is estimated to exceed $115 million in FY 1990 constant dollars for RDT&E, or $540 million for procurement in FY 1990 constant dollars, then it is categorized as a Major MTA (OUSD[A&S], 2019, p. 13). The use of the two separate dollar figures maps directly to the two distinct processes within MTA (OUSD[A&S], 2019, p. 13). That is, Rapid Prototyping is likely to weigh its estimated costs against the published RDT&E figure, as that is the most likely programmatic funding that would be applied in most circumstances (OUSD[A&S], 2019, p. 13; McGarry, 2022, pp. 20–21). Similarly, Rapid Fielding would most likely allocate procurement dollars to finance a fielding initiative (OUSD[A&S], 2019, p. 13; McGarry, 2022, pp. 20–21).

If the initial program cost estimate exceeds the established monetary thresholds and places the MTA program into the major system category, the component acquisition executive is responsible for furnishing additional traditional acquisition program documents, namely an approved requirements document, a cost estimate, a lifecycle sustainment plan (LCSP), and an abbreviated acquisition strategy or Simplified Acquisition Management Plan (SAMP) (GAO, 2023a, p. 14). Figure 4 illustrates these differences below.

Table 1. Documentation Requirements by System Type. Source: Department of the Air Force (2021), p. 13

	Major System[1]	Non-Major System[2]
Rapid Prototyping	• ADM signed by the DA • Approved Requirement[3] • Acquisition Strategy[4] • Cost Estimate[5]	• ADM signed by the DA
Rapid Fielding	• ADM signed by the DA • Approved Requirement[3] • Acquisition Strategy[5,6] • Cost Estimate[5] • Lifecycle sustainment plan[7]	• ADM signed by the DA

Additional oversight requirements include program information, data updates, and program objective memoranda (POM) (GAO, 2023a, p. 14). Such oversight ensures identification and collection of data to inform decision-making by program managers and DoD leadership (GAO, 2023a, p. 14).

F. RAPID PROTOTYPING AND RAPID FIELDING

Beyond the distinction between major and non-major MTA programs, there also exist two distinct processes for developing and fielding operational capabilities (OUSD[A&S], 2019, pp. 11-12). The first is MTA RP, which serves as a conduit for considering innovative technologies and new capabilities in rapid development efforts, resulting in fieldable prototypes that demonstrate new capabilities within five years (DAU, 2023, p. 46). Prototyping efforts fall broadly into two types: physical prototypes created through manufacturing processes, or virtual prototypes created using advanced modeling and simulation, augmented or virtual reality (OUSD[R&E], 2022, pp. 4-5). The capabilities provided by the prototype must be demonstrated in an operationally relevant environment and provide a residual level of capability that meets an emergent need (OUSD[R&E], 2022, pp. 17, 34-35). Rapid prototyping (RP) also requires that transition plans be provided within two years of program start (OUSD[A&S], 2019, p. 11). The early requirement for transition planning enables planning for a wide range of potential outcomes that may result from a rapid prototyping effort (OUSD[A&S], 2019, p. 11; OUSD[R&E], 2022, p. 21).

These outcomes range from discarding the prototype if it has limited utility after the project is completed to returning it to the technology base for further modification, testing, and evaluation (OUSD[R&E], 2022, pp. 19–20). A return for modification can occur when the prototype demonstrates operational utility but is neither immediately ready nor needed for current operations yet might be useful in future operational environments (OUSD[R&E], 2022, pp. 19–20). Transition into operational use is another possible scenario in which prototypes are deployed to address an immediate need, either as an interim solution until a program of record, usually via transition from RP to Rapid Fielding (RF) or MCA, or even as a final capability solution in its current form (OUSD[R&E], 2022, pp. 19–20). A transition to the MTA RF pipeline is warranted if the prototype meets the needs of the warfighter or reduces the life-cycle cost of a currently fielded weapon system, and the prototype has been demonstrated in an operationally relevant environment (OUSD[R&E], 2022, p. 20). Finally, a transition into an existing MCA program or transition into a new acquisition program utilizing one of the five materiel solutions pathways is also a possible outcome (OUSD[R&E], 2022, pp. 19-20).

The Rapid Fielding (MTA RF) pathway is intended for mature technologies that can be quickly moved into production and fielded as new or upgraded capabilities with minimal additional development (OUSD[A&S], 2019, p. 4). Under MTARF, programs are required to initiate production within six months of program start and complete fielding within five years (OUSD[A&S], 2019, p. 5). Rapid Fielding begins with a review of existing non-developmental items (NDIs) and proven technologies to satisfy the operational needs of the Joint Chiefs of Staff and Combatant Commanders (DAU, 2023, p. 67). The operational need is conveyed as an approved requirement and approved by the decision authority (DA) (DAU, 2023, p. 67). The requirement outlines a minimum fielding criteria and identifies the full funding required to complete the fielding (OUSD[A&S], 2019, p. 11). Once the approved requirement is received, the service responsible for the program develops a streamlined process that results in a succinct Service-level requirements document, approved by the Component Acquisition Executive (CAE) (OUSD[A&S], 2019, pp. 5,13).

Neither MTA pathway can be planned to exceed the five-year project execution timeline and is also subject to the program management triple constraint of cost, schedule, and performance (OUSD[A&S], 2019, pp. 4–5). Project start is defined as the date the acquisition decision memorandum (ADM) is signed by the DA. Some exceptions to this protocol exist; however, they require Defense Acquisition Executive (DAE) waiver authority (OUSD[A&S], 2019, pp. 4-5).

G. PROGRAM INITIATION

MTA program initiation is the CAE's decision and is based on validated operational need. Beyond validation, program managers have significant latitude to tailor documentation, artifacts, and even requirements for an MTA program (LaCamera, 2019, p. 20). MTA grants program managers significant latitude to incorporate only the documents, reviews, and assessments they need to support the acquisition strategy, which was tailored early in the program lifecycle to meet the specific characteristics, execution requirements, and risk profile of the specific effort (DAU, 2023, p. 32). Thus, program managers are empowered to tailor innovative, disciplined approaches to achieve success while still meeting the responsibilities of acquisition strategy development, program planning, fielding capabilities, and reporting program status, as required by the specifics of their program and applicable statutes (DAU, 2023, p. 33).

For approved MTA RP and MTA RF programs, the documentation protocol is intentionally designed to be less burdensome than that of an MCA acquisition process. It is generally exempt from DoDD 5000.01 documentation requirements (GAO, 2023, p.12). There will often be some variation in the total documentation package, depending on the program manager's situational understanding of programmatic needs (GAO, 2023, p. 13). The minimum required documentation for a major MTA program is outlined in Figure 5.

Table 2: DOD Directed Processes and Documentation Requirements for Middle Tier of Acquisition (MTA) Programs

Path	Required processes	Resulting documents
Rapid prototyping and rapid fielding MTA	Develop requirements	• Succinct requirements document within 6 months of MTA initiation.
	Implement acquisition and full funding strategies	• Acquisition strategy • Cost estimate
Rapid prototyping MTA	Consider innovative technologies and new capabilities to meet certain needs	• Acquisition decision memorandum at initiation that validates the use of the MTA pathway and identifies full funding required. Process also results in an approved requirement.
	Demonstrate performance and evaluate current operational purposes of proposed products and technologies	• Test strategy or an assessment of test results (included in the acquisition strategy)
	Transition successful prototypes to new or existing acquisition programs	• Transition plan (included in the acquisition strategy) providing a timeline for completion within 2 years of required documentation for transition
Rapid fielding MTA	Consider existing products and proven technologies to meet certain needs	• Acquisition decision memorandum at initiation that includes minimum fielding plan criteria and identifies full funding required. Process also results in an approved requirement.
	Demonstrate performance and evaluate current operational purposes of the proposed products and technologies	• Test strategy or an assessment of test results (included in the acquisition strategy)
	Consider life-cycle costs and address issues of logistics support, training, interoperability, and cooperative opportunities, among other things	• Life-cycle sustainment plan
	Identify and exploit opportunities to reduce total ownership costs	• Life-cycle sustainment plan
	Transition successful programs to operations and sustainment.	• Transition plan (included in the acquisition strategy) providing a timeline for completion within 2 years of required documentation for transition

Source: GAO analysis of Department of Defense (DOD) policy. | GAO-23-105008

Figure 5. Required Processes and Documentation Requirements for MTA. Source. Government Accountability Office (2023a), p. 12

Major MTA programs are required to formally execute the processes and produce the documentation listed in Figure 5. An important caveat not captured in the GAO table but situationally applicable to major MTA programs is the applicability of testing strategies and assessments related to cybersecurity, electromagnetic (EM) spectrum, chemical biological radiological and nuclear (CBRN), space, and interoperability requirements as applicable (OUSD[A&S], 2019, p. 11).

H. SERVICE-SPECIFIC NUANCE

DoDI 5000.80 discourages selecting the MTA pathway for developing major systems intended to fulfill requirements for significant joint force procurements or that involve considerable international partner involvement (DAU, 2023, p. 12). This prescription for a combination of flexibility at the programmatic level and the discouragement of programmatic efforts at the interagency level has spawned some Service-specific distinctions within MTA. The DoDI also mandates that each Service creates its own streamlined process that results in an executable requirements framework, and for its CAE to serve as the decision authority for their respective MTA efforts

(OUSD[A&S], 2019, p. 8-9). This is unique amongst the AAF pathways, and a defining departure from more traditional acquisition models.

1. Army Approach

The Army captures its MTA guidance in a policy memorandum issued by the Office of the Assistant Secretary of the Army for Acquisition, Logistics, and Technology (ASA[ALT]) (DOA[ASA(ALT)], 2020). Central to the Army's adaptation of DoDI 5000.80 guidance is centralized control of MTA efforts within the Army to the CAE (ASA(ALT)) (DOA[ASA(ALT)], 2020). The CAE must approve every Army decision to utilize the MTA pathway (DOA[ASA(ALT)], 2020, p. 3). Regarding Rapid Prototyping, the Army is the only Service that has mechanisms to approve (through the CAE) Rapid Prototyping initiatives before a validated requirements document is generated, "for rapid prototyping, the effort might precede a validated requirement, and in fact, may inform the requirement" (DOA[ASA(ALT)], 2018, p. 11). This statement could be interpreted as a full embrace of the DoD Prototyping Guidebook where the purpose of rapid prototyping is to generate information that supports a decision, and that success of a prototype is measured by whether it provides an actionable data set of information, and failure is more so defined by the quality level of the collected data versus what the prototype achieved or failed to achieve. (OUSD[R&E], 2022, pp. 2-3, 6).

2. Air Force Approach

The Air Force takes a somewhat different approach in its Service-level guidance. The Air Force's DAFI 63-146 serves as a direct supplement to DoDI 5000.80, adding Air Force-specific guidance directly into the DoDI, and republishing the combined document as DAFI 63-146 (DAF [SAF/AQ], 2021). This embedded approach is unique among the Services, which have all opted to issue stand-alone Service instructions. DAFI 63-146 provides program managers with guidance to develop acquisition strategies that include milestones, metrics, and execution guardrails, along with any other required documentation to the CAE, who is the Assistant Secretary of the Air Force for Acquisition, Technology, and Logistics (SAF/AQ) Staff (DAF [SAF/AQ], 2021, pp. 9, 14-17). The SAF/AQ provides oversight to MTA activities for the Air Force and delegates responsibilities to the Program Executive Office (PEO) or Milestone Decision Authority (MDA) as required

(DAF [SAF/AQ], 2021, p. 10). DAFI 63-146 further stipulates dedicated requirements documents for both Rapid Prototyping and Rapid Fielding pathways (DAF [SAF/AQ], 2021, p. 14, 17-18). These items include approved requirements documents, acquisition strategy, program information (goals, objectives, schedule, scope of decision reviews, planned follow on activities), contract strategy, funding strategy, risk management processes, guardrails against formal oversight breaches, test planning documents, IP strategies, and mandated oversight from the Air Force Deputy Chief of Staff (DAF [SAF/AQ], 2021, p 14-15).

The Air Force also designates Air Force Test and Evaluation (AF/TE) to act as the oversight body for test and evaluation of MTA capabilities and authorizes the use of existing testing and evaluation protocol to tailor testing and evaluation scope and methodology to best suit the MTA capability being developed (DAF [SAF/AQ], 2021, p 15).

3. Navy Approach

The Department of the Navy captured its guidance in SECNAVINST 5000.2G (Department of the Navy [DON], 2022). This instruction covers all six pathways within AAF and serves as the Navy's implementation of the Defense Acquisition System (Department of the Navy [DON], 2022, pp. 1-2). The instruction provides clear guidance for initiation of the MTA pathway, with the deliverables consisting of an approved requirement document, acquisition strategy, capability documentation, cost estimate, and LCSP (DON, 2022, Encl. 3, Table 4-1). Rapid Prototyping initiatives require a prototyping plan, acquisition plan, cost estimate, capability documentation, and a master test strategy (DON, 2022, Encl. 3, Table 4-2). Rapid Fielding enables the program manager to recommend to the decision authority which documents best support program execution. Outside of statutorily required documents (namely, FAR, DFARS, NMCARS-associated documents, and program initiation documents), everything else should be aimed at optimizing production and fielding timelines (DON, 2022, Encl. 3, Para. 6). This could be characterized as a full embrace of the tailoring mandate in DoDI 5000.02, which gives Naval RF initiatives significant latitude compared with the other Services.

4. USSOCOM Approach

United States Special Operations Command (USSOCOM) takes yet another nuanced approach to MTA operations under its purview. USSOCOM's policy is captured in USSOCOM Directive 71-4 (USSOCOM, 2020). This document serves as the basis for a Service-level capabilities and integration development system, known as SOFCIDS. The SOFCIDS process was created to fulfill the authorities granted by MTA in acquiring special operation peculiar capabilities (USSOCOM, 2020, p. 3). SOFCIDS also governs the use of Combat Evaluations (CV), a rapid assessment of system capability through direct application in combat environments, and is a unique USSOCOM provision (USSOCOM, 2020, p. 7). Rapid Prototyping within USSOCOM requires "capability sponsors" (i.e., USSOCOM leaders who believe they have a capability gap) to nominate their capability requirement via a special operations rapid requirements document (SORRD) (USSOCOM, 2020, p. 8). This SORRD at a minimum must contain the operational context, capability required, capability gap, and traceability, proposed program summary (i.e., cost, schedule, and materiel solution), threshold performance attributes to close the capability gap described in the capability gap and traceability section, and program affordability (USSOCOM, 2020, p. 8).

Rapid Fielding also requires capability sponsors to submit nominations using SORRD (USSOCOM, 2020, p. 8). The contents of this type of SORRD are not specifically outlined but instead should be commensurate with the additional risk associated with MTA RF, as it may impact multiple components, have fiscal implications well into the POM cycle, and may be subject to specific guidance contained in OUSD policy memoranda (USSOCOM, 2020, p. 8).

Upon receipt of the SORRD at either the component G8 or the sub-unified command, J8 assembles a rapid working group (RWG) to determine whether the SORRD nomination meets the MTA authority criteria or if it can be fulfilled via a more appropriate process. This is then forwarded as a recommendation by the SOF AT&L who is the CAE for USSOCOM (USSOCOM, 2020, pp. 8-9). This represents a much more operational, command-driven approach to MTA programming as compared to other services. It is sponsor-driven, presumably by operational-level commands within USSOCOM

(USSOCOM, 2020, pp. 8-9).

I. SERVICE DIFFERENCE

The removal of JROC-validated requirements from the MTA process can be seen as a significant break and even a reversal in the trend towards jointness that has been a feature of US military planning since the Goldwater-Nichols Act (OUSD[A&S], 2019, p. 5). This has indeed created significant divergence in published oversight of respective Service MTA efforts.

These differences reflect the Services, their unique cultures, and their requirements. For instance, the USSOCOM's SOFCIDS requirements underscore its inherently joint nature, and its replication of the joint requirements system within USSOCOM is somewhat unsurprising from this perspective (USSOCOM, 2020, pp. 6-7). The Army's lack of a defined requirement to start an MTA RP program (DOA[ASA(ALT)], 2018, p. 11) is potentially reflective of two things: the lower total costs of ground combat systems in comparison to the air and sea domains, and the near-instant evolution of ground-based tactics often spurned by low-cost solutions through history. Hedge row cutters repurposed from Nazi beach obstacles and installed on M4 Sherman tanks within a month of the initial D-Day landings in France (Blumenson, 1993, p. 206), or the proliferation of low-cost and often commercial drones that have proliferated in the battlespace in the Russo-Ukraine War (Fried, 2025) are just two examples.

There are also differences in philosophy for documentation. SECNAVINST 5000.2G leaves the minimum required documentation for a Rapid Fielding program at what is defined by statute, and anything beyond this threshold is a discretionary decision to the respective Service level decision authority (DON, 2022, Encl. 3. Table 4-2). This may reflect the more concrete nature of rapid fielding initiatives within the Navy and an interest in retaining maximum speed and flexibility in system fielding and in incremental improvements.

The Air Force, on the other hand, requires significantly more for the MTA RF pathway. These requirements go beyond the approved requirement, acquisition strategy, cost estimate, and lifecycle plan, and include additional contracting strategies, IP strategies, a test evaluation master plan (TEMP), risk management documentation, execution documentation that includes timing, scope, and decision reviews, follow-on activities, and relationship to other programs (DAF [SAF/AQ], 2021, pp. 14-18, 22). It could be interpreted that these additional documentation guardrails exist due to the Service's history of programmatic cost overruns in the post-milestone B MCA space, which is MTA RF's closest preexisting corollary, and a desire to proactively apply controls for future MTA RF acquisitions. Figure 6 provides post-milestone B cost growth in the F-35 program as an example.

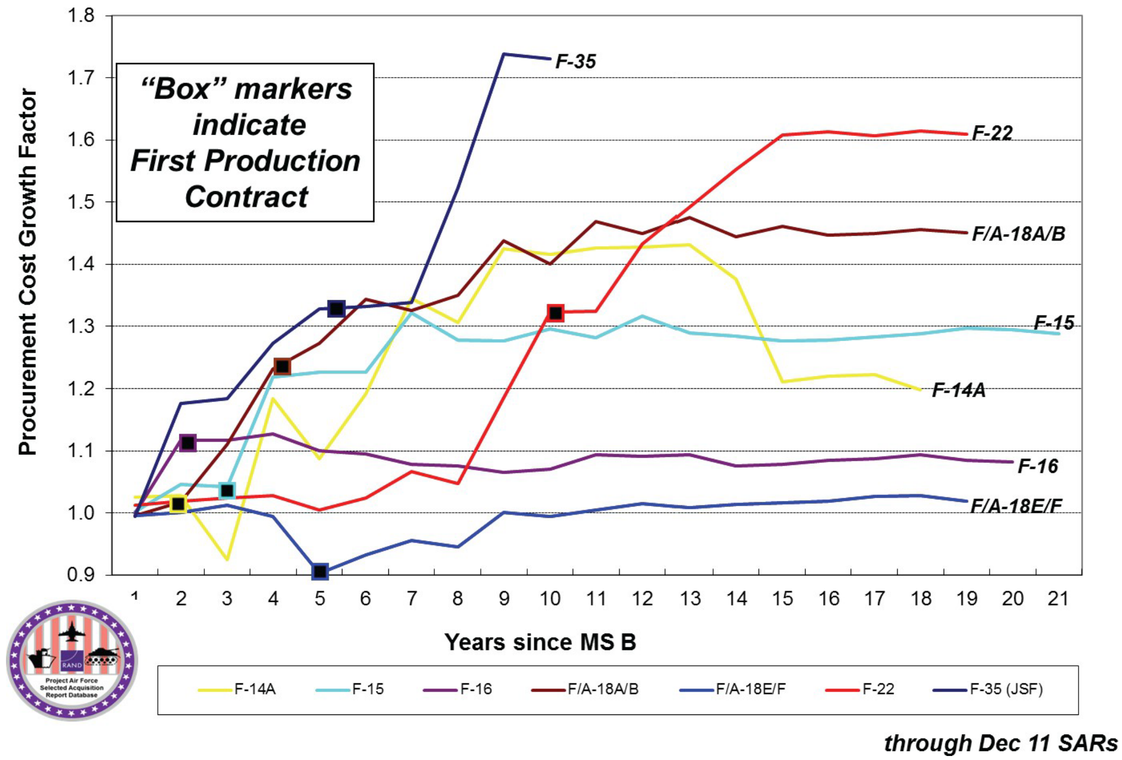

Figure 6. F-35 and Selected Fighter Post Milestone B Cost Growth. Source. Leonard & Wallace (2014), p. 4

The F-35 program is perhaps the best recent example of post-production decision cost growth in Air Force procurement and represents the largest such growth for a US fighter aircraft in history (Leonard & Wallace, 2014, p. 4).

J. PROGRAM TRANSITION

Beyond the stated intent to deliver capabilities to the warfighter in two to five years, another goal of many MTA-coded programs is to achieve a successful transition into the MCA pathway at the appropriate milestone (GAO, 2025a). Figure 8 graphically depicts a

Rapid Prototyping program into an MCA Milestone B and a Rapid Fielding program into a Milestone C.

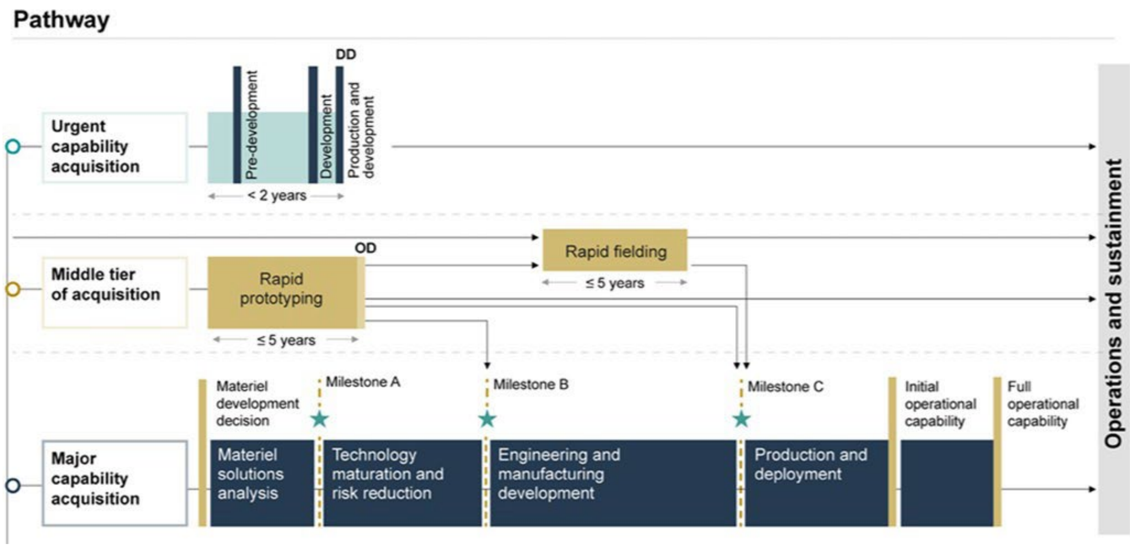

Figure 7. MTA Programmatic Transition. Source: Government Accountability Office (2025b), p. 7

This is a crucial step in sustaining the rapid momentum MTA is meant to drive within the WAS. Both require a transition plan to be submitted to the OUSD[A&S] (OUSD[A&S], 2019, pp. 11-12). The requirement provides a critical link to DoD oversight at the limit of the statutory authority granted under MTA and prepares many programs for Acquisition Category I (ACAT) I DoD-level oversight (OUSD[A&S], 2020d, pp. 9, 19-20).

K. CONCLUSION

This chapter establishes the statutory, policy, and organizational contexts to better understand MTA within the broader Warfighting Acquisition System. It sought to explain “big A” acquisition by examining its constituent parts, the JFRP, PPBE, and the AAF. The ways in which these systems interact with one another to shape the acquisition decisions, and how and where MTA fits into the AAF by providing an outlet to accelerate prototyping and fielding within a time-constrained framework. This chapter also discussed the differences between rapid prototyping and rapid fielding, major and non-major MTA efforts and their respective requirements, and the potential for program transition to other AAF pathways. Finally, it examined the Service-specific implementation protocols, which are unique to MTA alone within the larger AAF, as decision authority is held with the CAE and typically bypasses JFRP processes. This statutory, policy, and organizational background provides the foundation for the literature review, which will focus on prior research and oversight findings related to programmatic best-practice characteristics, MTA execution, and program outcomes.

THIS PAGE INTENTIONALLY LEFT BLANK

III. LITERATURE REVIEW

This literature review examines the statutory, policy, oversight, and academic foundations of the Middle Tier of Acquisition (MTA). It contextualizes how rapid acquisition authorities are intended to function and how they perform in practice. Together, these sources establish the analytical foundation and justify the use of a comparative, case-based methodology to examine how system characteristics and acquisition pathway selection influence MTA outcomes.

A. SECTION 804, FISCAL YEAR 2016 NDAA (MIDDLE TIER OF ACQUISITION)

Section 804 of the FY 2016 National Defense Authorization Act (NDAA) represents the legislative origin of the Middle Tier of Acquisition. It outlines Congress's intent to create a faster, more flexible acquisition category capable of delivering warfighting capabilities within five years (NDAA 2016, 2015). The statute directs the DoD to produce guidance for a "middle tier" separate from urgent needs processes and the conventional major capabilities acquisition framework, to be completed within two to five years (NDAA 2016, 2015).

Section 804 defines two acquisition pathways: rapid prototyping and rapid fielding (NDAA 2016, 2015). The rapid prototyping pathway emphasizes the accelerated development and demonstration of emerging technologies with the goal of delivering a usable capability within a five-year period following program approval (NDAA 2016, 2015). The rapid fielding pathway is intended for technologies that are already mature, requiring programs to transition quickly into production shortly after approval and achieve operational deployment within five years (NDAA 2016, 2015). Together, these statutory timelines demonstrate congressional intent to institutionalize speed as a central characteristic of the MTA framework (NDAA 2016, 2015).

In addition, Section 804 establishes the requirement for a streamlined, coordinated budgeting and acquisition process (NDAA 2016, 2015). The statute directs the Department to establish a process to produce an approved requirement within six months (NDAA 2016, 2015). It explicitly states that MTA programs are not subject to JCIDS or DoDI 5000.01, except where specifically allowed (NDAA 2016, 2015). This exemption represents a deliberate effort to reduce bureaucratic steps within the acquisition process.

Section 804 also outlines implementation expectations for both rapid prototyping and rapid fielding efforts, including the competitive evaluation of candidate technologies, methods for assessing prototype effectiveness under realistic operational conditions, and mechanisms for integrating successful prototype efforts into follow-on programs. (NDAA 2016, 2015). These provisions demonstrate Congress's recognition that many promising technologies stall before reaching a program of record and that transition mechanisms must be built into the pathway from the outset.

To enable execution, the statute authorizes streamlined management structures, including direct reporting chains for program managers, annual performance evaluations tied to program outcomes, and authorities to make trade-offs among requirements, cost, and schedule (NDAA 2016, 2015). It also establishes a Rapid Prototyping Fund to support early prototyping efforts, thereby designating a dedicated fund for MTA and the rapid prototyping pathway (NDAA 2016, 2015).

B. GOVERNMENT ACCOUNTABILITY OFFICE LITERATURE ON MIDDLE TIER OF ACQUISITION

Government Accountability Office (GAO) assessments provide a central foundation for evaluating the effectiveness of the MTA pathway because, unlike policy or doctrinal publications, GAO reports provide independent, program-level evidence regarding technical maturity, schedule execution, pathway compliance, and organizational behavior across rapid prototyping and rapid fielding efforts. This body of literature is particularly relevant to research examining why MTA outcomes vary substantially across programs and why statutory flexibility alone does not guarantee accelerated capability delivery.

GAO's 2023 review of MTA oversight identifies systemic limitations in the DoD's governance and monitoring of MTA programs (GAO, 2023a). Although DoDI 5000.80 requires standardized reporting of pathway selection, schedule objectives, and technology maturity, GAO finds that inconsistent reporting practices and incomplete data fields significantly constrain portfolio-level visibility (GAO, 2023a). Program offices frequently misreport or omit core attributes, including pathway designation, operational demonstration timelines, and technology readiness levels (TRLs) (GAO, 2023a). In many cases, existing reporting systems do not allow programs to update key information after restructuring or transition between rapid prototyping and rapid fielding pathways (GAO, 2023a). As a result, senior acquisition leaders lack the decision-quality data required to distinguish between stand-alone MTA programs, increments embedded within larger acquisition efforts, and hybrid structures that combine multiple pathways (GAO, 2021).

From a governance perspective, the GAO concludes that MTA's flexibility increases, rather than reduces, the importance of formal oversight mechanisms (GAO, 2023a). Without standardized reporting of system maturity, program scope, and pathway intent, acquisition leaders cannot identify early indicators of misalignment between system characteristics and pathway selection (GAO, 2023a). This deficiency weakens the component's ability to intervene before technical risk, schedule compression, and unrealistic transition planning erodes the intended advantages of rapid acquisition authorities (GAO, 2023a).

Across GAO's 2020–2025 Weapon Systems Annual Assessments, the MTA pathway is consistently characterized as a speed-focused mechanism (rapid prototyping / rapid fielding) meant to deliver operational capability within a compressed timeline, but GAO repeatedly finds that execution and oversight practices often undermine those objectives (GAO, 2020b). GAO reports portfolio-level visibility problems with MTA programs that lack comparable milestones and use inconsistent schedule metrics, limiting OSD's ability to aggregate performance and track cost/ schedule outcomes consistently across the enterprise (GAO, 2020b, p. 64). GAO also identifies weaknesses during program initiation, noting that many MTA efforts proceed without foundational business-case documentation, including acquisition planning artifacts, reliable cost projects, and formal assessments of schedule or technology risk (GAO, 2022). These weaknesses during program initiation limit the leadership's ability to evaluate whether programs can realistically achieve statutory and policy-drive timelines (GAO, 2022). As a result, GAO documents recurring delays and technology maturation gaps that increase technical risk during transition and can necessitate additional development beyond the MTA window (GAO 2023b). Finally, GAO highlights that many MTA programs plan to transition into major capability acquisition, often at development, concluding that MTA efforts frequently function as a front-end to long, linear acquisition timelines, in some cases implying overall delivery horizons exceeding 15 years when MTA initiation does not reduce total cycle time (GAO, 2023b, 2024).

GAO attributes these extended timelines to structural features of program design, particularly insufficient technology maturity at program entry (GAO, 2025). Programs attempting to mature multiple critical technologies within the limited duration of MTA face compounded schedule risk, as TRL advancement historically requires multi-year development cycles even under stable funding management conditions (GAO, 2025). Consequently, MTA programs that begin with immature technologies tend to replicate the same sequential development patterns characteristic of traditional acquisition pathways, thereby neutralizing the intended speed advantage of the pathway (GAO, 2025).

GAO also identifies systematic performance differences across system categories. Hardware-centric programs, vehicle upgrades, and integration-focused platforms demonstrate comparatively stronger alignment with MTA constraints, while software-intensive and human-system integration-heavy programs experience higher rates of redesign, soldier usability challenges, and pathway resets (GAO, 2024). These trends suggest that system complexity, degree of user interaction, and dependency on iterative software refinement materially influence whether programs can achieve accelerated delivery under MTA authorities (GAO, 2024).

In parallel, GAO critiques the continued reliance on linear development models within MTA programs (GAO, 2024). Although many programs report adopting modular open systems approaches, few consistently employ minimum viable product strategies, digital twins, or integrated digital threads within a unified iterative development framework (GAO, 2025). GAO emphasizes that these tools are most effective when implemented collectively to support rapid experimentation, progressive capability releases, and continuous refinement, rather than as isolated technical enhancements embedded within traditional sequential development structures (GAO, 2025).

GAO's analysis of the Adaptive Acquisition Framework further demonstrates that iterative development principles remain unevenly institutionalized across acquisition pathways (GAO, 2025a). The software acquisition pathway incorporates structured feedback loops, frequent user evaluations, and MVP-based delivery. MTA and MCA pathways lack comparable formal mechanisms (GAO, 2025a). As a result, many programs interpret "flexibility" as reduced documentation or streamlined approvals rather than as an operational mandate for continuous learning and design iteration (GAO, 2025a, p. 29).

GAO concludes that procedural acceleration alone is insufficient to generate sustained schedule compression for complex cyber-physical systems (GAO, 2024). Instead, acquisition speed depends on structural features such as recurring user feedback, incremental capability releases, and deliberate design-test-refine cycles (GAO, 2024). Without these elements, digital engineering tools and alternative contracting strategies are absorbed into fundamentally linear program architectures, limiting their impact on cost, schedule, and performance outcomes (GAO, 2024).

Collectively, GAO's literature supports several propositions that underpin this study's comparative case study. First, MTA outcomes are path-dependent and are primarily shaped by system maturity, integration complexity, and organizational execution choices rather than by statutory authority alone (GAO, 2023a). Second, low technology readiness at program entry is a dominant predictor of schedule extension and pathway transition (GAO, 2025). Third, hardware-dominant systems align more consistently with MTA constraints than software-intensive or human-centered systems (GAO, 2024). Fourth, oversight deficiencies obscure early warning indicators of program risk (GAO, 2023a). Finally, true acquisition speed is driven by iterative development structures rather than by deductions in documentation or formal milestone requirements (GAO, 2025a).

C. RAND CORPORATION LITERATURE ON MIDDLE TIER OF ACQUISITION

RAND Corporation research provides a system-level analytical foundation for evaluating why the MTA pathway produces uneven outcomes across programs. Unlike program-by-program oversight studies, RAND's work focuses on acquisition policy design, data governance, performance measurement, and decision behavior within the Adaptive Acquisition Framework (AAF). Three reports are particularly relevant to this study: *Building a Broader Evidence Base for Defense Acquisition Policymaking* (Bartels et al, 2020), *Using Metrics to Understand the Performance of the Adaptive Acquisition Framework* (McKernan et al., 2022), and *Issues with Access to Acquisition Information in the Department of Defense* (Drezner et al., 2020). Collectively, these studies support the thesis that MTA success depends on alignment among system characteristics, pathway design, and institutional data practices rather than on statutory authority alone.

Bartels et al. (2020) argue that acquisition policymaking has traditionally relied on empirical historical program data to evaluate reforms, but this approach is poorly suited to assessing new acquisition constructs such as MTA. Their analysis identifies four structural assumptions underlying data-driven policy evaluation: that historical data are generalized to new policy contexts, that future operating conditions will resemble the past, that causal relationships between policy and outcomes can be isolated from other variables, and that the risks of policy experimentation are acceptable (Bartels et al., 2020, p. 2-3). The authors conclude that these assumptions do not hold in many cases for non-incremental reforms, including the introduction of new acquisition pathways (Bartels et al., 2020, p. 2-3).

As a result of MTA's alterations to governance, documentation requirements, and schedule constraints, historical data from Major Capability Acquisition (MCA) programs provide limited predictive value for how programs will behave under rapid prototyping or rapid fielding authorities (Bartels et al., 2020, p. 2-3). RAND therefore cautions that observed program outcomes alone may obscure underlying structural misalignment between system maturity, integration complexity, and pathway selection (Bartels et al., 2020, p. 2-3).

To address this limitation, the authors introduce policy gaming as a complementary analytical tool for evaluating acquisition reforms (Bartels et al., 2020, p. 4-5). In a prototype acquisition policy game centered on the MTA, RAND researchers simulated pathway-selection decisions for hypothetical programs that varied in cost, schedule pressure, technology maturation, system complexity, and warfighter demand (Bartels et al., 2020, p. 4-5). Participants representing program managers, contracting officers, technical experts, and operational users consistently weighted technology readiness and integration complexity more heavily than urgency when determining whether programs were suitable for MTA execution (Bartels et. al., 2020, p. 4-5). Programs characterized by immature technologies or complex integration requirements were viewed as high risk for schedule and performance failure under accelerated timelines (Bartels et al., 2020, p. 4-5.

McKernan et al. (2022) extend RAND's system-level perspective by developing a formal metric framework for evaluating pathway performance under the AAF (McKernan et al. 2022, p. 3-4). Using a logic model methodology that links mission objectives, pathway goals, intended outcomes, and observable attributes, the authors propose standardized portfolio-level metrics to assess whether each acquisition pathway is achieving its intended purpose (McKernan et al., 2022, p. 4).

For the MTA pathway, the study identifies five core indicators of pathway health: (1) average percentage cost growth, (2) schedule deviation between program starts and operational demonstration or initial operational capability, (3) technology readiness level at program entry, (4) budget stability over time, and (5) the number of prototypes fielded, transitioned, or terminated (McKernan et al., 2022, p. 7). These measures are designed to capture not only speed but technical feasibility and program survivability within compressed timelines (McKernan et al., 2022, p. 7).

McKernan et. al. (2022) identifies several challenges unique to MTA performance assessment. First, most MTA efforts are non-major programs for which only limited standardized data is collected, constraining portfolio-level analysis (McKernan et al., 2022, p. 5-6). Second, the pathway contains heterogeneous program types, rapid prototypes, and rapid fielding efforts, making direct comparison difficult (McKernan et al., 2022, p. 5-6). Third, leadership's reluctance to impose additional reporting requirements in the name of speed reduces the availability of consistent data for oversight and learning (McKernan et al., 2022, p. 5-6).

McKernan et al. (2022) further observe a continual tension between MTA's schedule constraints and the information requirements needed for effective governance (McKernan et al., 2022, p. 5). This tension mirrors the challenges identified in the government oversight literature and contributes to the Department's limited ability to distinguish between pathway design flaws and program-specific execution problems (McKernan et al., 2022, p. 5). The authors recommend longitudinal and case-based analyses using standardized metrics to identify recurring success factors and failure modes

(McKernan et al., 2022, p 8-9).

Drezner et al. (2020) examine acquisition data governance in the post-2016 statutory environment and provide the most detailed treatment of how MTA complicates enterprise-level information management. The study identifies acquisition data as foundational to program management, portfolio oversight, and system-wide learning, yet finds that recent organizational restructuring and pathway diversification have fragmented data ownership and weakened standardization (Drezner et al., 2020, p. 1-2).

Drezner et al. (2020) argue not that MTA programs are intentionally designed with austere reporting requirements to preserve speed, but this design choice has produced unintended consequences: inconsistent data definitions, limited visibility into technical maturing and schedule realism, and difficulty aggregating program information across the Department (Drezner et al., 2020, p. 9-10).

Drezner et al. (2020) identify three structural challenges created by MTA. First, most MTA programs are lower-dollar-value efforts that fall outside traditional reporting regimes, leaving oversight organizations dependent on ad hoc data calls (Drezner et al., 2020, p. 9-10). Second, the pathway's mixed composition of prototypes and fielded systems complicates the establishment of uniform data standards (Drezner et al., 2020, p. 9-10). Third, the removal or downgrading of legacy reporting mechanisms, such as the Selected Acquisition Report and the Defense Acquisition Executive Summary, has disrupted continuity in longitudinal program data, further constraining comparative analysis across pathways (Drezner et al., 2020, p. 19-22).

RAND recommends that acquisition data requirements be driven by decision use cases rather than by historical reporting conventions (Drezner et al., 2020, p. 29-30). Drezner et al. (2020) advocate for an enterprise-wide acquisition data strategy, alignment of pathway reporting with standardized data frameworks, and selective retention of legacy reporting elements to preserve analytical continuity. Importantly, Drezner et al. (2020) caution that rapid-acquisition authorities cannot substitute for institutional learning mechanisms, and that speed without visibility risks the repeated reinvention of known failure patterns.

D. ETEMANDI AND KAMP: MIDDLE TIER ACQUISITIONS AND INNOVATION

Dr. Amir Etemadi and Dr. John Kamp (2022a) examined the use of MTA authorities to accelerate system innovation and, by extension, assessed whether MTAs were used as intended. In their literature review, the researchers conduct a comparative analysis of MTA program execution with selected MCA pathways and examined 53 programs, comprising 12 MTAs and 41 MCAs, selected from the 2020 GAO-selected Acquisition Report (Etemadi & Kamp, 2022a, p. 118).

They further divide the research into three distinct intervals or phases: program approval through Milestone B (St-MS. B), Milestone B through critical design review (MS. B-CDR), and critical design review through initial operational capability (CDR-IOC) (Etemadi & Kamp, 2022a, p. 118). This is done to more precisely identify schedule acceleration within the programmatic dataset, with acceleration used as a proxy for process innovation (Etemadi & Kamp, 2022a, p. 118-119). They also code these programs as being either modular, agile, agile-modular, or "no code" meaning the program did not exhibit traits of either (Etemadi & Kamp, 2022a, p. 118).

Their statistical analysis of these subcategories reveals that MTA programs are consistently faster on key timeline metrics, specifically in approval to Milestone B (St-MS. B) and in critical design review to initial operational capability (CDR-IOC) (Etemadi & Kamp, 2022a, p. 119-120). These findings align closely with the design logic for MTA programs, as most programs operate in these two spaces: St-MS. B represents MTARP initiatives, and CDR-IOC represents MTARF activities (Etemadi & Kamp, 2022a, p. 119-120). Thus, the time compression observed by Etemadi and Kamp suggests that MTA is working as a designed accelerator within the discrete program set examined (Etemadi & Kamp, 2022a, p. 119-120).

The analysis of programs coded for agile, modular, agile & modular, or none of the properties revealed a strong alignment between agile and shorter schedule durations (Etemadi & Kamp, 2022a, p. 120-121).

E. ETEMADI AND KAMP: ESTIMATING MIDDLE TIER ACQUISITION SCHEDULE RISK

Etemadi and Kamp further model and assess schedule performance of MTA programs when compared directly to MCA programs, specifically within the Air and Missile commodity types, from the 2020 GAO annual weapon system assessment (Etemadi & Kamp, 2022b, pp. 492-493, 500). The researchers collect 27 unique programs for analysis and conduct a Monte Carlo analysis to estimate probable schedule outcomes derived from the collected data (Etemadi & Kamp, 2022b, p. 492).

The simulation results demonstrate that MTA programs were more likely than not to deliver intended capabilities within their 60-month schedule limitation (Etemadi & Kamp, 2022b, p. 498-499). The simulations reveal key program attributes that achieved the schedule advantage: a short time to contract award, either through commercial-type contracts or by reducing contractual requirements, and an intentional reduction in program complexity and technical risk (Etemadi & Kamp, 2022b, p. 499). In terms of programmatic schedule segments, the most differentiated between MTA and MCA was the interval between critical design review and initial operational capability (CDR-IOC) (Etemadi & Kamp, 2022b, p. 498-500). The simulations demonstrated that, on average, MCA efforts take almost four times as long to proceed from CDR to IOC, suggesting that MTA systems must be less complex to produce and field in comparison to MCA (Etemadi & Kamp, 2022b, p. 499-500). MTAs also carry less schedule risk, in absolute terms, than MCAs when adjusted for schedule duration (Etemadi & Kamp, 2022b, p. 499).

F. PERDUE AND TEISSONNIERE-ALMODOVAR: EMPIRICAL ASSESSMENT OF MIDDLE TIER OF ACQUISITION EFFECTIVENESS

Perdue and Teissonniere-Almodovar (2024) conducted a comprehensive academic evaluation of the effectiveness of the Middle Tier of Acquisition (MTA) pathway. They examine whether the MTA rapid prototyping pathway achieves its statutory objective of delivering capabilities at the "speed of relevance," with particular attention to measurable program outcomes and available acquisition data sources (Perdue & Teissonniere-Almodovar, 2024, p. 1-3).

Perdue and Teissonniere-Almodovar (2024) limited their analysis to rapid prototyping efforts due to the relative immaturity of the pathway and the scarcity of completed rapid fielding programs (Perdue & Teissonniere-Almodovar, 2024, p. 1-3). Effectiveness was operationalized as whether a program transitioned to a program of record, restructured into another acquisition pathway, or was terminated (Perdue & Teissonniere-Almodovar, 2024, p. 5-6). Using DAVE data, the study conducted hypothesis testing across a sample of 55 completed MTA rapid prototyping programs (Perdue & Teissonniere-Almodovar, 2024, p. 53-54). The analysis estimated that between 71.2% and 92.2% of programs transitioned or restructured following prototyping, suggesting a high nominal rate of pathway "success" when measured strictly by continuation outcomes (Perdue & Teissonniere-Almodovar, 2024, pp. 27–29).

To contextualize these results, the Perdue and Teissonniere-Almodovar (2024) attempted to correlate DAVE program records with MTA programs referenced in GAO oversight reports to identify common predictors of success or failure (Perdue & Teissonniere-Almodovar, 2024, p. 58-60). This effort was unsuccessful due to inconsistent program identifiers, incomplete technical and schedule data, and limited descriptive detail within the DAVE dataset (Perdue & Teissonniere-Almodovar, 2024, p. 58-60). As a result, the study concluded that existing acquisition databases are insufficient to support robust causal analysis of MTA effectiveness at the portfolio level (Perdue & Teissonniere-Almodovar, 2024, pp. 30–32).

Perdue and Teissonniere-Almodovar (2024), therefore, argue that quantitative program-level outcome measures alone cannot account for variation in performance across MTA efforts (Perdue & Teissonniere-Almodovar, 2024, p. 30-32). They identify the lack of standardized reporting on technology maturity, system integration burden, requirements stability, and program risk posture as major obstacles to evaluating whether the pathway itself or the characteristics of the systems entering it drive observed outcomes (Perdue & Teissonniere-Almodovar, 2024, p. 30-32). The authors explicitly recommend that future research adopt a comparative case study design to examine similar programs with divergent outcomes and identify structural success factors and failure modes (Perdue & Teissonniere-Almodovar, 2024, p. 73-75).

This conclusion directly reinforces concerns raised in government and policy literature regarding the trade-off between acquisition speed and analytical visibility under MTA authorities. By demonstrating that high transition rates coexist with weak explanatory power, the study highlights the risk of conflating administrative progression with operational or technical success. The Perdue and Teissonniere-Almodovar (2024) further note that classification restrictions and nondisclosure agreements governing DAVE access significantly constrain external validation and replication of acquisition research, limiting the transparency necessary for institutional learning (Perdue & Teissonniere-Almodovar, 2024, p. 73-75).

From a methodological perspective, Perdue and Teissonniere-Almodovar's (2024) quantitative analysis of DAVE-based MTA outcomes establishes an important baseline while simultaneously exposing the limitations of outcome-only metrics for evaluating rapid acquisition pathways (Perdue & Teissonniere-Almodovar, 2024, p. 27-32). Its findings suggest that effectiveness assessments must move beyond binary continuation measures to account for system-level attributes such as technological readiness, software dependence, integration complexity, and user-system interaction, factors not systematically captured in current databases.

G. BUB, KAITLYN: ANALYSIS OF RAPID PROTOTYPING WITH THE DOD

Bub (2023) examined the execution realities of rapid prototyping programs conducted under the MTA pathway, with emphasis on how the statutory 5-year schedule constraints reshape program risk management, cost control, and performance trade-offs (Bub, 2023, p. 1-3). Conducted as a Naval Postgraduate School capstone project, the study focused solely on rapid prototyping efforts, given the greater availability of completed programs and publicly released data relative to rapid fielding initiatives (Bub, 2023, p. 3-5).

Bub (2023) ground the analysis in project management theory, specifically the "triple constraint" model that links cost, schedule, and performance (Bub, 2023, pp. 12–15). Bub argues that the fixed-schedule requirements embedded in the MTA alter this relationship by removing schedule as a flexible variable. As a result, cost growth, technical performance degradation, or reduced scope become the primary mechanisms through which programs absorb risk when development challenges arise (Bub, 2023, p. 14-16). This structural condition, rather than program mismanagement alone, is identified as a recurring driver of downstream transition difficulties (Bub, 2023, p. 16-18).

Methodologically, Bub (2023) employed a qualitative literature review supplemented by analysis of program metrics derived from government reports, including those from the GAO, RAND, Defense Acquisition University (DAU), the Inspector General, and other public sources (Bub, 2023, p. 6-8). Access restrictions associated with DAVE limited the release of detailed program-level technical data, thereby constraining the ability to conduct more in-depth quantitative modeling (Bub, 2023, p. 8-10). Consequently, the study emphasized root-cause analysis to identify recurring barriers to successful transition from prototype to program of record or field capability (Bub, 2023, p. 10-12).

Bub (2023) identified several persistent execution challenges across MTA rapid prototyping efforts: insufficient technology maturity at program entry, optimistic assumptions regarding integration feasibility, compressed testing timelines, unstable or evolving requirements, and limited opportunity for iterative user feedback (Bub, 2023, pp. 21-25). Bub further observed that while MTA accelerates initial development, it does not eliminate the underlying technical work required to mature complex systems, often displacing that effort into later phases or follow-on acquisition pathways (Bub, 2023, p. 23-25).

Bub (2023) cautions against interpreting transition rates as a sufficient indicator of success. Programs may formally transition while carrying unresolved technical debt, incomplete logistics planning, or immature production processes, increasing the probability of cost growth and schedule instability in subsequent acquisition phases (Bub, 2023, pp. 26–28). These findings parallel to broader concerns in the oversight literature regarding the risk of equating administrative progression with operational readiness (Bub, 2023, p. 36-38).

H. WEBSTER: THE MIDDLE TIER OF ACQUISITION: SPEED VS RIGOR

Webster (2024) provided a policy-centered examination of the MTA pathway in a DAU-sponsored research paper. Webster (2024) situates MTA within four decades of acquisition reform efforts and evaluates whether the pathway meaningfully advances the DoD's objective of delivering military capability at the "speed of relevance" (Webster, 2024, p. 1-2).

Webster (2024) employed a qualitative methodology grounded in thematic analysis of statutory sources, policy documents, government oversight reports, academic literature, and professional commentary spanning the period from the 1986 Packard Commission through the end of 2023 (Webster, 2024, p. 7-10). Webster organized the literature around three central questions: why the MTA authority was established, what objectives it was intended to serve, and whether current implementation practices are achieving those objectives (Webster, 2024, p. 5-6).

Webster (2024) traced the emergence of MTA to converging pressures from strategic competition, rapid commercial technology cycles, and long-standing dissatisfaction with the cost and duration of traditional major defense acquisition programs (Webster, 2024, p. 13-21). The analysis emphasized that MTA represents not merely procedural streamlining but a deliberate institutional trade-off in which documentation requirements, milestone reviews, and formal reporting mechanisms were reduced to accelerate program initiation and execution (Webster, 2024, p. 23-27).

Webster (2024) identified a body of literature regarding MTA effectiveness. Advocates within the acquisition community highlight improved responsiveness, increased use of commercial technologies, and shortened development timelines. Critics, including multiple oversight organizations, emphasize diminished cost transparency, inconsistent technical baselines, weakened portfolio visibility, and the erosion of standardized performance reporting (Webster, 2024, p. 49-52). Webster concluded that both perspectives are partially valid and reflect the structural tension inherent in rapid acquisition authorities (Webster, 2024, p. 52-55).

A central contribution of Webster (2024) is its articulation of "speed versus rigor" as a persistent policy dilemma rather than a temporary implementation flaw (Webster, 2024, pp. 12-16). Webster argued that accelerating acquisition processes necessarily constrains the quantity and quality of information available for oversight and long-term learning. Without deliberate institutional mechanisms to capture data on cost, schedule, technical maturity, and operational effectiveness, MTA risks reproducing historical acquisition failures under compressed timelines (Webster, 2024, p. 12-16).

Webster (2024) further cautioned against using MTA as a substitute for disciplined acquisition planning or as a mechanism to bypass major capability acquisition oversight thresholds. Instead, Webster recommended selective integration of MTA programs with other Adaptive Acquisition Framework pathways, improved cost estimation practices, and the development of standardized performance metrics to preserve strategic transparency without negating schedule advantages (Webster, 2024, pp. 18–20).

I. LIMITATIONS OF LONG-CONTEXT DATA PROCESSING IN LANGUAGE MODELS

Liu et al. (2023) examine the effectiveness of large language models (LLMs) in processing extended input context, specifically assessing whether increased context length translates into improved information utilization. Their study is particularly relevant to this research, as it directly informs the reliability and validity of LLM-assisted qualitative and quantitative data analysis methods (Liu et al., 2023, p. 1-2).

Liu et al. (2023) employ controlled experimental designs across two primary tasks, multi-document question answering and key-value retrieval, to evaluate how models identify and use relevant information embedded within long input sequences (Liu et al., 2023, p. 2-3). Their findings indicate that model performance is highly sensitive to the positional placement of relevant information within the input context. Lie et. al (2023) observed a U-shaped retrieval and reasoning pattern in LLM performance, where models demonstrated stronger accuracy when prominent information appeared near contextual boundaries and weaker performance when critical details were located within the interior of long-context inputs (Liu et al., 2023, p. 1-2).

This phenomenon has important implications for research methodologies that rely on LLMs to synthesize large volumes of textual data, such as document coding, thematic extraction, or comparative case analysis. In relation to this study, where structured document analysis and aggregation of acquisition program data are central to the methodology, the findings suggest that LLM outputs may not uniformly reflect all input data. Instead, there exists a potential bias toward information positioned at the boundaries of the dataset, which could influence analytical conclusions if not properly mitigated (Liu et al., 2023, p. 2-4).

Additionally, Liu et al. (2023) demonstrate that increasing the volume of contextual information does not necessarily improve model performance. In some cases, the inclusion of additional documents introduces noise and reduces accuracy, indicating a trade-off between context breadth and effective information retrieval (Liu et al., 2023, p. 3-5).

The study further highlights that even advanced or extended-context models do not consistently outperform their standard counterparts in utilizing long inputs, reinforcing the conclusion that current LLM architecture remain limited in their ability to fully leverage expanded context windows (Liu et al., 2023, p. 2-3). This results in methodological rigor in data structuring, such as prioritizing key information, segmenting inputs, or employing iterative querying, is essential when integrating LLMs into research workflows.

Lie et al. (2023) provides critical empirical evidence that challenges assumptions regarding the effectiveness of long-context processing in LLMs. Their findings underscore the need for deliberate input design and cautious interpretation of outputs when applying these tools to complex analytical tasks.

J. OPERATIONALIZING BUSINESS CASE WITHIN DEFENSE ACQUISITION

The GAO's Cost Estimating and Assessment Guide provides a useful foundation for the operationalization of a business case as a measurable acquisition construct and more than an opaque management term. The GAO highlights the need for reliable cost estimates to support funding decisions, resource evaluations at critical decision points, and performance measurement baselines to empirically judge program performance and improve resource allocation to increase the probability of program success (GAO, 2020a).

The guide provides a concrete example of how the GAO operationalizes its understanding of a sound business case through two specific case studies. These were Case Study 15: Independent Cost Estimates, from Coast Guard Acquisitions, and Case Study 17: Incomplete Risk and Uncertainty Analysis, from Coast Guard Acquisitions (GAO, 2020a, pp. 125, 141). These two case studies looked at a past Coast Guard attempt to acquire three icebreaker vessels (GAO, 2020a, p. 141). The Coast Guard failed to establish a comprehensive cost estimate that included the entirety of the proposed vessels' lifecycles and instead was baselined on variations in costs only occurring during the development and production phases of the program (GAO, 2020a, p. 141). It excluded any risk and uncertainty analysis around the operations, support, and disposal phases of the system lifecycle (GAO, 2020a, p.141). This baseline was established before the Preliminary Design Review (PDR) in early post-milestone B (GAO, 2020a, p. 323). In effect, locking in the poorly constructed lifecycle cost estimate, and creating the potential for significant program instability if design and schedule estimates experienced any drift or deviation during the Engineering and Manufacturing Development (EMD) phase (GAO, 2020a, p. 323). The resulting full lifecycle cost estimate of $9.8 billion was ultimately not a credible estimate as it lacked sufficient assessment of total lifecycle costs and ultimately established an unrealistic baseline (GAO, 2020a, p. 324). It is evident that, from the GAO's perspective, the business case construct must adequately articulate whether a program's acquisition strategy is grounded in sufficient knowledge to make credible commitments regarding cost, schedule, performance, and technical risk.

The concept of business case is also important to systems engineering literature and frames acquisition decisions as structured evaluations of alternatives, and effective decisions consider the full system lifecycle, which includes development, operation, support, and disposal to ensure that proposed solutions are feasible and sustainable over the life of the system (INCOSE, 2015, pp. 25-27). This is most effectively accomplished through a structured decision-management framework that converts a broad decision problem into a recommended course of action and a corresponding implementation plan (INCOSE, 2015, p. 110). The International Council on Systems Engineering (INCOSE) handbook emphasizes that a significant portion of total lifecycle costs is committed in the earlier stages of program evolution (INCOSE, 2015, p. 14). Demonstrating the significance of a sound business case informing early programmatic decisions is a key risk management tool for programmatic cost and schedule.

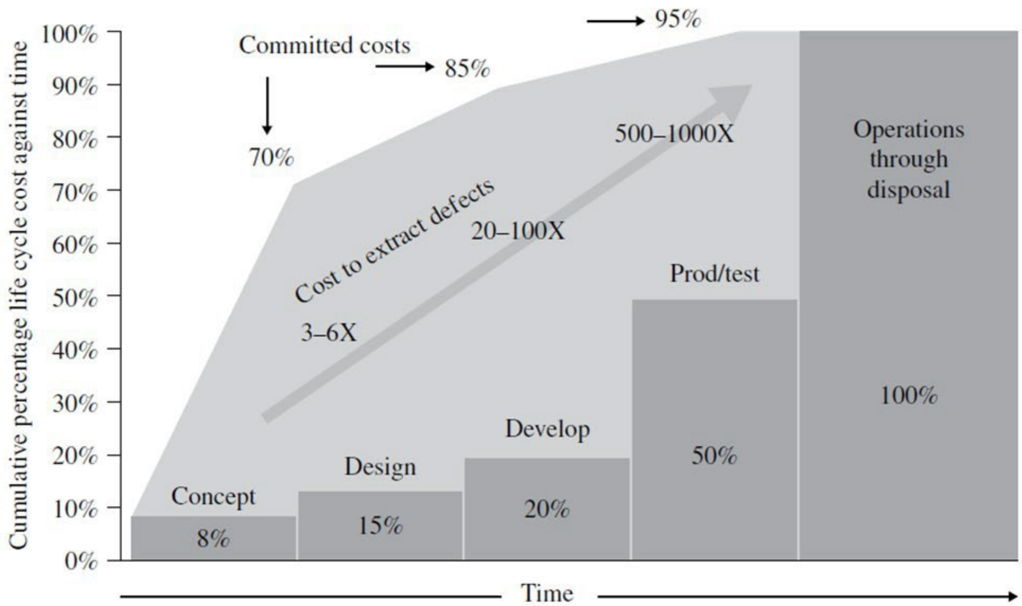

Figure 8. Committed Lifecycle Costs Over Time. Source: International Council on Systems Engineering (2025), p. 14

From an industry perspective, the importance of system architecture and its relationship to a program's business case strength is present. The National Defense Industrial Association suggests that effective implementation of Modular Open Systems Approaches (MOSA) and the well-constructed system definition, stable interfaces, as well as the ability to integrate and upgrade modules over time, are important factors in enabling cost-effective and executable lifecycle sustainment protocols that materially enhance a program's business case (NDIA, 2023, pp. 3-4, 7-8).

K. TECHNOLOGICAL READINESS AS A RISK AND FEASIBILITY CONSTRUCT

The concept of technological readiness is another essential construct in defense acquisition because it enables evidence-based assessment of the critical technologies required to support system development and integration (OUSD[R&E], 2025, p. viii). A Technology Readiness Assessment (TRA) evaluates the maturity of critical technology elements (CTEs), which are defined as technologies essential to system performance that are new technologies/processes or novel approaches, or that pose significant technical challenges (OUSD[R&E], 2025, pp. 28-29). An assessment of these CTEs is translated into Technological Readiness Levels (TRLs) and is used as the standardized metric for technological maturity within defense acquisition programs (OUSD[R&E], 2025, pp. 34-35).

TRAs ultimately serve as significantly consequential decision-support tools for program offices and provide evaluations of technologies and their demonstrated capabilities in relevant environments (OUSD[R&E], 2025, p. 34). Thereby, informed decisions can be made about the feasibility of integrating the technology in question and the associated technical risks (OUSD[R&E], 2025, p. 1). A composite of programs' TRAs forms an Integrated Technical Risk Assessment (ITRA) (OUSD[R&E], 2025, p. 1). This assessment is then used to support informed trades among cost, schedule, and performance, and to inform the Milestone Decision Authority's (MDA) approval of program milestones or consideration of revisions to program requirements (OUSD[R&E], 2025, pp. 1, 23).

Figure 9. Five-Step Process for Conducting High Quality TRA. Source: Office of the Under Secretary of Defense for Research and Engineering (2025), p. 25

The maturity of the technologies in these assessments has been identified as a significant determinant of acquisition outcomes (OUSD[R&E], 2025, p. viii). The GAO found that programs incorporating immature technologies are significantly more likely to experience cost growth and schedule delays, and that there is a negative relationship to these factors for programs with mature technologies prior to integration (GAO, 2016, pp. 26-29).

The relationship between technical maturity and MTA is uniquely critical, as these programs must progress rapidly to meet their compressed timelines and deliver operationally relevant capabilities (OUSD[A&S], 2019). TRL is particularly important within the MTA context because it serves as the estimate of whether critical technologies are mature enough to support program execution within MTA's statutorily compressed timeline (DoD, 2019, p. 4; OUSD[R&E], 2025, pp. 34–35). Programs with lower overall

TRLs might require additional maturation, testing, and integration efforts, increasing the risk that a program might not deliver its capability within the fixed schedule constraint (GAO, 2016, pp. 26-29).

L. IMPLICATIONS OF ITERATIVE DEVELOPMENT

The Information Technology Acquisition Advisory Council (IT-AAC) and the Foundation for the Defense of Democracies (FDD) in response to Executive Order 14265, entitled "Modernizing Defense Acquisitions and Spurring Innovation in the Defense Industrial Base" stated that in order to transform defense acquisition, empowerment and reform of the defense acquisition workforce requires embedding innovative metrics within the higher level restructuring completed in efforts such as the adoption of AAF (IT-AAC and FDD, 2025, p. 4). Among the process improvements IT-AAC recommended to achieve meaningful modernization are a transition to a mission-driven acquisition model aligned with agile commercial practices and the widespread adoption of open architecture principles, and abandoning the Department of Defense Architecture Framework (DoDAF) entirely (IT-AAC and FDD, 2025, p. 4).

Anton et al. (2020) suggest that MTA Rapid Prototyping may be complementary to iterative development practices as a methodology for delivering more responsive capability sets to the warfighter. Their analysis suggests that iterative approaches are applicable when requirements are more stable, when there is trade space among cost, schedule, and performance, and when partial solutions are acceptable during development (Anton et al., 2020, pp. 23-25). This logic aligns with DoD prototyping guidance, which frames prototyping to generate information to support decisions, reduce technical and integration risk, validate designs, gain warfighter feedback, and inform achievable requirements (OUSD[R&E], 2022, pp. 1–4). Programmatic efforts with significant development or integration challenges can benefit from iterative approaches, especially those whose requirements are more flexible, incorporate user feedback, and utilize incremental refinement of system architectures (Anton et al., 2020, pp. 48-49, 115). This uncertain but flexible environment allows programs to "fail fast" and identify technical limitations, validate assumptions, and reduce overall uncertainty through experimentation and feedback (OUSD[R&E], 2022, pp. 3–4). RAND's analysis supports viewing MTA RP and iterative development as complementary strategies in accelerating acquisition programs (Anton et al., 2020). Although they are not fully interchangeable concepts, MTA RP provides the pathway and informs decision making, while iterative approaches support learning, feedback, and system refinement in uncertain and somewhat flexible environments (Anton et al., 2020, pp. 20-21, 118-119).

M. DELEGATION OF AUTHORITIES WITHIN ORGANIZATIONAL STRUCTURES

The level of delegation for decision-making authorities within organizational hierarchies can be a point of contention and is often shaped by competing demands between responsiveness and control within a given organization. McElheran found that delegation is strongly associated with high adaptation value, defined as how much a decision requires a localized understanding of conditions, and that local actors have a localized information advantage (McElheran, 2012 pp. 10-11). Centralized decision-making, however, is associated with coordination requirements, particularly where system-wide integration, standardization, or interoperability are required (McElheran, 2012, pp. 11-12). The opposite scenario occurs where delegation is preferred if the perceived adaptation benefits are greater than the costs of coordination, and centralization when the benefits of coordination across the team outweigh the value derived from a localized information advantage (McElheran, 2012, p. 7-9). This descriptive delegation framework may lend credence to a more decentralized approach, especially for MTA RP endeavors, where prototyping would not necessarily require extensive coordination at the CAE or OSD level (OUSD[A&S], 2019, pp.13-14). Similarly, lower-level decision-makers may be better suited to make decisions in which there is technological uncertainty, a need to rapidly incorporate user feedback, and a need to refine designs within a statutorily compressed timeline.

N. MODULAR OPEN SYSTEMS ARCHITECTURE AND DEFENSE APPLICATIONS

The National Defense Industrial Association (NDIA) published “Modular Open Systems Approach: Implementation Challenges and Opportunities,” in which it discusses the ongoing integration of MOSA approaches into the wider defense acquisition space.

MOSA has emerged as a central construct in modern defense acquisitions, which are often characterized by complexity, evolving requirements, and the need for sustained technological capabilities throughout the system lifecycle (NDIA, 2023, pp. 1-2). MOSA provides a structured approach to managing system architectures, interfaces, and integration within the system-of-systems type environments (NDIA, 2023, pp. 3-5). Effective MOSA implementation depends on the alignment of system and subsystem hierarchies, well-defined interfaces, and architectural boundaries that enable the introduction of modular components at the appropriate level within a given system and support future technologies through this modularity (NDIA, 2023, pp. 3-4). This is often achieved through structured Work Breakdown Structures (WBS) and reference architecture to clearly define hierarchies and subsystem boundaries, and MIL-STD-881 serves to formalize the defense acquisition use of WBS (NDIA, 2023, p. 4). MIL-STD-196 defines levels of taxonomy and modularity; a system taxonomy breakdown assists with the appropriate structure, level, and naming conventions needed for successful MOSA integration (NDIA, 2023, p. 5). This standardization, configuration, and interface management regime allows for clearly defined stakeholder responsibilities and can help reduce the risk of interface divergence throughout a system's lifecycle (NDIA, 2023, p. 5). Thoughtful implementation of MOSA principles is critical within MTA programs, where timelines are unforgiving for poorly defined interfaces and architectural boundaries that increase integration risk and, as a result, programmatic risk, as time to resolve incompatibilities could well exceed statutorily defined limits (NDIA, 2023, pp. 3-5, OUSD[A&S], 2019, p). The promise of modularity through MOSA supports faster adaptation by reducing integration tasks to smaller, cohesive modules rather than re-integrating entire system architectures to accommodate emergent requirements or upgraded technologies (NDIA, 2023, pp. 7-8, OUSD[A&S], 2022a, p. 13).

O. CONCLUSION

The existing body of literature on MTA provides a foundation for understanding the intent, structure, and observed performance of MTA RP and MTA RF pathways. GAO oversight reports consistently identify systemic challenges related to TRL, reporting inconsistencies, and limited portfolio-level visibility (GAO 2024). RAND studies extend this understanding by highlighting deficiencies in acquisition data governance and the difficulty of evaluating pathway performance using traditional empirical approaches (Drezner et al., 2020). Academic research further contributes by demonstrating that MTA programs can achieve schedule acceleration under certain circumstances, particularly when technical complexity is constrained and agile or modular development approaches are employed (Etemadi & Kamp, 2022a).

Despite these contributions, the literature reveals several limitations. First, much of the existing research relies on aggregate-level analysis or incomplete datasets, constraining the ability to identify causal relationships between program characteristics and outcomes (Perdue & Teissonniere-Almodovar, 2024). Second, current studies predominantly employe outcome-based metrics, such as transition rates or schedule performance, which do not fully explain why some programs succeed within MTA constraints while other experience delays or require pathway transition (GAO, 2023b). Third, there is limited research that compares similar programs with divergent outcomes, restricting insight into how factors such as technology maturity, system integration complexity, and execution approach influence pathway effectiveness (Perdue & Teissonniere-Almodovar, 2024).

As a result, while the literature establishes that MTA performance varies across programs, it does not fully explain the conditions under which the pathway is most effective or most likely to fail. This gap limits the ability of acquisition professionals to make informed pathway selection decisions and align program design with structural constraints of rapid acquisition authorities (GAO, 2024). This Capstone contributes to the broader acquisition literature by providing a structured framework for evaluating pathway-program alignment, offering insights that can guide program-level decision-making.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. METHODOLOGY

This research assesses the extent to which MTA RP programs reflect acquisition best practice contracts in their program artifacts contained within the Acquisition Information Repository (AIR) of the Defense Acquisition Visibility Environment (DAVE). The identified acquisition programs and all associated documentation available within AIR were then systematically evaluated using a rubric-constrained model that leveraged the recently adopted GenAI.mil Gemini Large Language Model (LLM) to identify, extract, and score best practices.

These outputs were then cataloged and used to analyze averaged scores between the best practices at the Service level, prevalence of each best practice within different categories of program documentation, and finally to assess the strengths of association between different best practices across all collected program data.

This methodology in part builds on Perdue and Teissonniere-Almodovar's (2024) assessment of MTA RP program outcomes. Perdue and Teissonniere-Almodovar highlighted two specific areas of interest for future research. First, they suggested that future research could “Analyze data reporting and collection methods to identify any patterns or trends that may affect the oversight and performance of the MTA pathway” (p. 80). Second, they proposed the following:

> Investigate the possibilities of utilizing artificial intelligence to aid DAVE in quantifying its reporting requirements. Researchers may be able to leverage large language models to investigate varies gaps within DAVE and provide levels of reporting for every system. This insight should highlight issues with compliance and should issue warnings to stakeholders to ensure higher quality when navigating DAVE (p. 81).

Rather than using LLMs to identify gaps in DAVE data, this research leverages their ability to systematically sort, analyze, and characterize the large existing datasets within DAVE’s AIR. This research also used GAO reference materials that are substantially overlapping to build the best practice framework and populate LLM prompts. The LLM assists in quantifying the extensive data library available within DAVE’s AIR into discrete ordinal measures. This meta-analysis of programmatic documentation seeks to uncover insights beyond purely outcome-level characterizations of successful and unsuccessful programs.

A. BEST PRACTICE IDENTIFICATION

The identification of best-practice constructs used in this study is informed by prior oversight and analytical work conducted by GAO, as well as other supporting DoD source material (e.g. DODIG, DoD Prototyping Guidebook) to identify best practices and leading development principles in defense acquisition. As outlined in the literature review, an emphasis on a recurring set of best practice areas is found throughout the literature. These best practices and principles were synthesized and articulated as:

1. Implementation of a MOSA Approach
2. Implementation of Iterative Development
3. Use of Modeling and Simulation Design Concepts
4. Delegation of Authority to the Lowest Level
5. Attainment of a Sound Business Case
6. Technology Readiness Level at Program Entry

Together, these areas represent interrelated dimensions of program execution, governance, and risk management that substantially define how MTA acquisition programs are structured and executed. The research is centered on MTA rapid prototyping programs, rather than rapid fielding efforts, which typically enter the acquisition process with greater tolerance for immature technologies and manufacturing readiness, resulting in comparatively less emphasis on TRL and MRL thresholds at program initiation for MTA RF endeavors.

The research analyzed DAVE AIR MTA program documentation for observable indicators of these principal constructs. Rather than attempting to measure program success through outcome metrics alone, this study evaluates the extent to which these best practices are reflected in program documentation and treats documented construct evidence as a proxy for programmatic intent and implementation.

This approach is driven by practical limitations, specifically involving the lack of direct observation of program execution, access to program personnel, and validated outcome data for the full lifecycle of MTA RP program efforts. As a result, the program information contained in AIR represents the most consistent and comprehensive set of artifacts available for comparison and analysis.

B. DATA SELECTION AND ACQUISITION INFORMATION REPOSITORY

The research applied a definitional distinction between successful and unsuccessful based on the program status filter applied to all programs within AIR. Specifically, coded successful programs as those having been labeled as the following within AIR.

- Transitioned/Restructured—Implies that the program has moved out of the MTA RP pathway and into another AAF pathway for further development and refinement. Along with this transition comes a significant restructuring of a program's scope, schedule, and funding to align with the selected new pathway.
- Residual Capability Delivered—Signifies that a prototype or minimally viable product has been delivered to the warfighter for operational use.
- Full Operational Capability (FOC) / Full Deployment (FD)—Indicates the program has reached its intended user and that user is now fully equipped and trained in the employment of the system. Full Deployment has a slightly different meaning; the total number of planned systems has been delivered to the end user.

Unsuccessful program labels and their definitions are described below:

- Terminated—Has been officially ended by the Milestone Decision Authority (MDA). This can happen for a number of reasons, such as poor performance, reaching cost or schedule thresholds, or lack of continued relevance in an evolving operational environment. All acquisition efforts cease upon termination.
- Disposal—The prototyping effort has concluded, the necessary data has been gathered, and the prototyping effort has not been slated for fielding, further use, or transition. While a valid and "successful" outcome of a prototyping project, it does not lead to production of an enduring equipment capability or transition.
- Suspended—Programmatic work has ceased. It may be restarted later. However, given the statutory time constraints on MTA efforts, suspension seriously jeopardizes, if not outright, leads to, the future cancellation or termination due to inadequate schedule space.

Using these pre-existing filters within AIR the research identified 84 successful MTARP programs and 21 unsuccessful MTARP programs. These programs exhibit wide variability in the completeness and number of documents included in the programmatic listing. To ensure completeness and reasonable inclusion in the analysis, programs' documentation had to include at least one of the following: Acquisition Strategy, Capabilities Development Document (CDD), Lifecycle Sustainment Plan (LCSP), Systems

Engineering Plan (SEP), or Test and Evaluation Master Plan (TEMP). These documents were selected as the minimum criterion because they were assumed to be the most critical acquisition planning documents.

Given the differences in documentation minimum requirements between major and non-major programs. Many of the successful and unsuccessful non-major programs were not included because they did not meet the minimum criteria. Ultimately, 39 successful MTARP programs or 46.4% of the programs within the AIR database, and 6 unsuccessful programs or 28.6% of programs within AIR containing the applicable status labels. Of the 39 successful programs selected, 25 (64.1%) were defined as major systems, and 14 (35.8%) were non-major systems. In terms of unsuccessful programs, only 1 of the 6 unsuccessful programs was a major system (16.7%), and the remaining 5 were non-major (83.3%). This filtering for documentation completion significantly improves data quality, but significantly reduces the sample size, and may limit the generalizability of the results to wider defense acquisition applications.

C. INQUIRY AND ANALYSIS USING A RUBRIC-CONSTRAINED LLM FRAMEWORK

Inquiry into this extensive data set, consisting of 356 unique program documents from 45 successful and unsuccessful MTA Rapid Prototyping programs, employed a structured prompt and rubric to provide an analytical decomposition of the identified best practice areas. This approach enables a consistent application of evaluation criteria across a large document library that would otherwise be impractical to assess manually.

These prompt templates were each geared towards the six identified best-practice approaches and sought an interpretation of the following:

- MOSA Prompt—An evaluation of the architectural openness of a given program. Indicators include concepts like open interfaces, modular design approaches, and competition-enabled incremental upgrades.
- Iterative Development Prompt—Identify and measure iterative development practices. Indicators include concepts such as agile methodology, feedback mechanisms, a minimum viable product release, and a strategy for delivering capabilities over time.
- Modeling and Simulation Prompt—Quantify the extent to which Modeling and Simulation concepts are embedded in program documentation. Indicators include concepts like digital twins, virtual environments,

hardware in the loop (HWIL), and software in the loop (SWIL).

- Delegation of Authority Prompt—Evaluate the depth and distribution of decision-making authority within a given acquisition program. This is quantified primarily by identifying where decision-making authority is held for various functions within program documents.
- Business Case Prompt—Quantify the strength and continued surveillance of a program's business case. Indicators include references to analysis of alternatives (AoA), cost targets, trade-offs and trade space, and the linkage of technical maturation to funding.
- Technology Readiness Prompt—Quantify the transparency and depth of program-level technological maturity assessments and the validation/ revalidation of critical technologies. Indicators include the assignment of TRL(s), the identification of critical technologies, and associated readiness and maturation plans.

Standardized prompts were chosen to constrain the model to evaluate only explicit documentary language, to avoid inference, to require direct quotation of supporting text, and to provide page references where available. The full prompt templates for each of the best practice criteria are in Appendix X.

D. SCORING METHOD

Each of the 356 unique program documents was attached and analyzed by GenAI.mil's Gemini-based LLM using a single prompt. This process was then repeated for each prompt. For each of the best practice constructs the LLM was instructed to:

1. Perform a full-document scan to identify all qualifying evidence.
2. Extract verbatim excerpts for every instance that corresponds to the construct.
3. Assign a single overall ordinal score between 0 and 3 to produce a summary measure of construct presence within each document.
4. Provide instance-level justification, explaining the rationale for linkage to the rubric criteria.

The ordinal scoring scale served as the primary mechanism for evaluating the inclusion and maturity of the best-practice concepts. The scoring scale was defined as follows:

- 0—Absent: No evidence of the construct.
- 1—Superficial: High-level declarative references without supporting mechanisms.
- 2—Intent and Partial Mechanism: Explicit intent with partial implementation detail.

- 3—Operationalized: Concrete and actionable mechanisms demonstrating full implementation intent.

The use of a rubric-based scoring system for LLM examination created a replicable system of qualitative and quantitative analysis that is constrained to verbatim text and tied to a standardized rubric and a bounded ordinal scoring construct. Each prompt enabled the operationalization of unique best-practice constructs across all applicable MTARP documents in AIR, resulting in a standardized measurement model for these qualitative constructs.

E. ARTIFACT CODING

After receiving the output from GenAI.mil Gemini LLM for each artifact and for each best practice. The outputs were recorded as a unique observation in a structured spreadsheet with columns for the following:

- Program Identifier
- Pathway (Rapid Prototyping or Rapid Fielding)
- Artifact Type (CDD, TEMP, etc)
- Best Practice Score (0-3)
- Supporting Text Extracts
- Section/page Reference

This structure enables aggregation across programs, document types, and constructs for subsequent analysis, and was used as the starting point for the data analysis.

F. DATA ANALYSIS

The analysis employs a multi-level analytical approach. The unit of analysis varies between methods to align with the underlying research questions. First, a heatmap analysis is employed at the document type (e.g. Acquisition Decision Memorandum, Transition Plan, etc.) level across all programs to both identify where the best practice constructs are concentrated within the program document ecosystem, as well as to evaluate levels of best practice variability between both successful/unsuccessful programs, and between document types. The second method is a Service-level aggregate of best-practice variability to examine differences between organizations. Finally, a linear regression is performed at the program level. This is done to examine potential relationships among various best-practice constructs.

This combined approach reflects a deliberate trade-off. Higher-level aggregation is necessary to interpret the structure of and detect patterns between the services and the types of commonly used program documents. While individual program-level analysis allows for analysis of variance needed to test relational assumptions. Collectively, these levels of analysis provide a complementary view of acquisition best practice maturity across the institutional, organizational, and programmatic levels.

G. BEST PRACTICE APPLICATION ACROSS ACQUISITION DOCUMENT TYPES

Heatmap analysis of the data used ordinal scoring to indicate best practice construct maturity. For each construct, ordinal scores (0-3) assigned at the individual document level were aggregated by calculating the mean across all relevant observations. Values closer to 3 indicate higher levels of construct operationalization, while values closer to 0 indicate minimal evidence. The heatmap itself is illustrated on a sliding scale, where red is equal to 0, yellow is equal to 1.5, and green is equal to 3; shading between these colors indicates a transitional state that does not exactly align with the three measures. Table 2 illustrates successful programs, and Table 3 is for unsuccessful programs.

Table 2. Successful Rapid Prototyping Program Heatmap

Document Types Across All Programs	N	MOSA AVG	Iterative AVG	M&S AVG	Delegation AVG	Business Case AVG	TRL AVG
Acquisition Decision Memorandum (ADM)	128	0.76	1.53	0.36	2.11	2.51	1.47
Acquisition Strategy	53	2.72	2.78	2.17	2.71	2.89	1.88
Transition Plan	28	1.39	1.70	0.70	2.00	2.54	1.38
Requirements Docs (ICD+CDD+CPD)	26	2.31	2.00	1.54	1.79	2.46	1.48
Cost Estimate	26	1.19	1.54	0.67	1.00	2.32	1.04
Lifecyle Sustainment Plan (LCSP)	11	2.36	2.27	2.10	2.00	2.91	1.64
Assessment of Test Results	10	2.00	2.33	1.22	1.11	2.40	2.25
Approved Requirement	8	0.13	0.75	0.00	1.50	1.00	0.78
Test and Evaluation Master Plan (TEMP)	4	3.00	3.00	3.00	2.00	3.00	2.75
APB	3	1.33	2.00	0.00	2.00	2.67	1.00
Program Assessment Report (PAR)	2	0.00	2.00	1.33	1.00	3.00	1.00
Program Notification	1	0.00	1.00	0.00	3.00	1.00	0.00
Program Protection Plan	1	3.00	3.00	1.00	2.00	3.00	1.00
Selected Acquisition Report (SAR)	1	2.00	2.00	0.00	1.00	2.00	1.00
Systems Engineering Plan (SEP)	1	3.00	3.00	2.00	2.00	3.00	2.00

Table 3. Unsuccessful Rapid Prototyping Program Heatmap

Doc Type Across Programs	N	MOSA AVG	Iterative AVG	M&S AVG	Delegation AVG	Business Case AVG	TRL AVG
Acquisition Decision Memorandum (ADM)	15	0.73	1.27	0.38	2.56	2.80	0.67
Acquisition Strategy	11	2.00	2.36	1.80	2.50	2.64	1.09
Cost Estimate	5	2.20	2.80	1.80	1.40	3.00	1.40
Transition Plan	5	1.80	1.80	0.83	2.17	2.50	1.17
Approved Requirement	4	1.00	4.00	0.50	1.50	2.25	1.20
Assessment of Test Results	4	1.75	1.25	2.00	0.00	1.50	0.75
Requirements Docs (ICD+CDD)	3	1.67	2.00	1.00	1.50	3.00	1.00
Program Termination Memo	2	2.00	2.50	0.00	3.00	3.00	1.00
Lifecyle Sustainment Plan (LCSP)	1	3.00	3.00	0.00	2.00	3.00	1.00
Test and Evaluation Master Plan (TEMP)	1	0.00	0.00	0.00	0.00	0.00	0.00
Cyber Security Strategy	1	2.00	2.00	0.00	2.00	3.00	2.00

The heatmap analysis provides insight into how MTARP best-practice constructs are operationalized across acquisition documents. At a high level of analysis, there appears to be no systemic superiority of successful program documentation over that of unsuccessful programs. Many document types and construct cells are roughly comparable between the two groups. Some document/construct pairings with sample sizes greater than 10 have equal or higher values in the unsuccessful programs. Delegation and business case metrics for Acquisition Decision Memoranda (ADMs) were higher for unsuccessful projects than those for successful counterparts. Conversely, successful program acquisition strategies were higher amongst MOSA and iterative development constructs using the same $n \geq 10$ threshold.

While there is no uniform observable higher mean amongst successful programs. Successful programs do tend to stand out in the Modeling and Simulation and Technological Readiness Level constructs. On average, across all document types, M&S constructs have higher mean scores, although the means for both successful and unsuccessful programs are uniformly the lowest scored of all constructs. Suggesting that successful programs may emphasize M&S more at the margins, but overall, the construct remains weakly applied across all sampled MTARP programs, regardless of outcome.

The technological readiness of successful programs also generates a higher mean score across comparable document types. This observation is perhaps the most intriguing of the heatmap analysis and suggests greater consideration for technological maturity, critical technologies associated with the programmatic effort. It is also worth noting that M&S and TRL are both technically grounded constructs, with TRL in the risk management domain and M&S in the execution domain. These two differences may indicate that the program outcomes are more closely associated with the depth of technical validation and the degree of risk reduction than with the other constructs examined.

Unfortunately, the sampling distribution of unsuccessful programs also does not readily lend itself to generalizable comparisons. There were only two document types (Acquisition Decision Memorandum and Acquisition Strategy) with sample sizes ≥ 10. The sample size asymmetry between successful programs (39) and unsuccessful programs (6) makes the unsuccessful program heatmap inherently more sensitive to individual programs and is potentially subject to significant skew. This makes it difficult to draw definitive conclusions between the two sample sets.

H. BEST PRACTICE APPLICATION AMONGST THE SERVICES AND PROGRAMMATIC OUTCOMES

At the Service level, best practice constructs were aggregated across programs in which a given Service was designated as the lead. Using the same mean ordinal scoring approach applied in the heatmap analysis, this aggregation enables the development of Service-level depictions of best-practice profiles through radar charts to visualize the dataset (displayed in figures 10 and 11).

1. Successful Service Level Program Comparison

Across the successful programs, the radar chart profiles from the Army, Navy, Air Force, and Space Force are broadly similar in their shape and magnitudes (Figure 10). All Services demonstrate comparable levels of iterative development and delegation of authority. There is minor variation within the MOSA construct, where the Space Force has a marginally higher mean score than do other Services, which are tightly grouped together at a lower mean score. Additionally, there is some variation in TRL and Business case, although it again is marginal in significance. The Army had the highest mean score for TRL and was somewhat higher than the other service means, as for the business case, both the Space Force and Army had higher mean averages than did the Navy and Airforce. Finally, within M&S, the Army was marginally better, with some separate from the next highest mean belonging to the Navy.

These overall similarities suggest that, within the successful program subset, there is no compelling evidence that any Service takes a significantly divergent approach to incorporating the best practice framework into their respective program documentation regimes. Instead, these programs appear to broadly share similar structures in how these constructs are included within program documentation.

Figure 10. Successful Programs by Service

2. Unsuccessful Service Level Program Comparison

The radar chart for unsuccessful programs presents a more limited comparison, as only the Army and Air Force had unsuccessful programs that met the inclusion criteria (Figure 11). Within this subset, there is somewhat more variability; however, these differences must be considered against the highly sensitive nature of this six-program data set, and the significant impact that single-program characteristics can have on this small sample. Ultimately, meaningful Service-level inferences cannot be extrapolated, and caution is required when interpreting the results. There appears to be greater emphasis from the Army on inclusion of the business case and modeling construct, and within M&S. There also is an observable difference in Air Force emphasis on MOSA, when compared to the Army within the six-program dataset.

The absence of additional Service data should not be interpreted as having not possessed unsuccessful MTARP programs, but rather a lack of minimally viable program files. Within the 21 unsuccessful programs in AIR, all services were represented (Army, Air Force, Navy, Space Force); however, only six programs from the Army (2) and the Air Force (4) contained minimally viable documentation to be included in the larger meta-analysis.

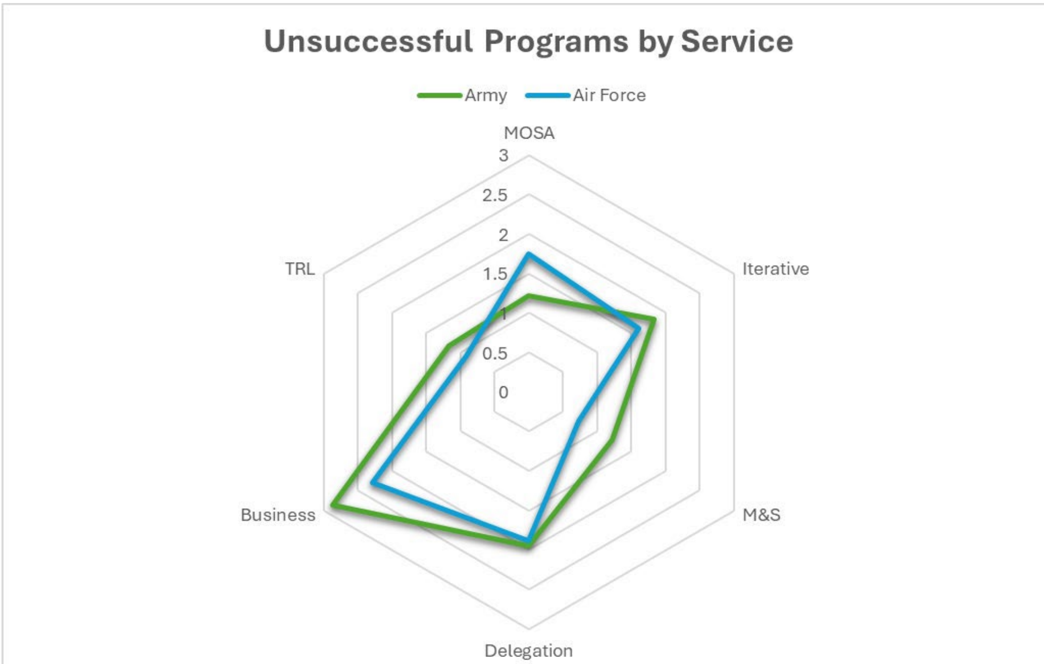

Figure 11. Unsuccessful Programs by Service

I. SUCCESSFUL VS. UNSUCCESSFUL COMPARISON

A comparative analysis of successful and unsuccessful programs provides a direct evaluation of the influence that the identified best practices might have on program outcomes. However, and again, this comparison must be interpreted with due consideration and caution given the substantial imbalance in sample sizes, and the overall low sample size (6) for unsuccessful MTA RP programs.

Within both successful and unsuccessful program types, similar mean values were recorded for MOSA, delegation of authority, and business case constructs. The only consistent differences appear in M&S and TRL (Figure 12). In the case of M&S, unsuccessful programs surprisingly demonstrate a marginally higher score, indicating greater relative strength in this area. Successful programs, on the other hand, had meaningfully higher TRL scores, with a 0.568 difference between the two mean values.

Cumulatively, the radar chart analysis reveals that the primary observable differences are limited to themes around program execution and risk reduction. Specifically, M&S and TRL where successful programs do exhibit marginally higher mean scores. This suggests that program outcomes may be more closely associated with an emphasis on technical readiness and analytical validation. The following section uses linear regression to explore this possibility as well as other possible relationships between the best practice constructs using successful program data set.

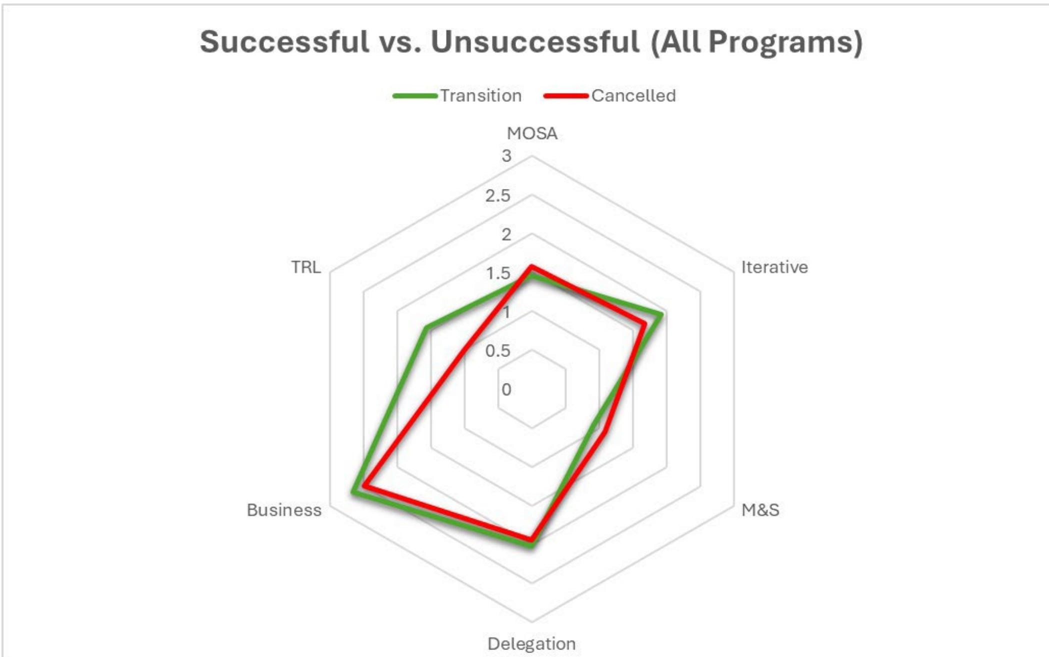

Figure 12. Successful vs. Unsuccessful

J. TESTING BEST PRACTICE RELATIONSHIPS USING LINEAR REGRESSION

Using linear regression to conduct exploratory inference on the construct variables. To complete methodically sound linear regressions, these constructs were decomposed into distinct analytical categories based on the functions they fulfill within the larger program management effort. The first category identified was the execution-integration domain, which encompassed the variables of iterative development, MOSA, and M&S. The relationship between these concepts was of interest primarily to gauge whether these development approaches were mutually reinforcing and co-implemented within programs, or whether they existed more as unique, independently implemented constructs.

The second area of analysis is contained within risk management and decision-making structures, personified in the business case and TRL constructs. The relationship between these variables was examined to assess the level of technical maturity and soundness of programmatic rationale are or are not integrated into other constructs, and by extension the coherence of risk management and decision frameworks within the data set.

The final area of analysis is within the realm of governance and focused on the role of delegated authority. Delegation was always treated as the independent variable to determine if increased decision authority at lower levels is associated with differences in the execution-integration and risk management and decision-making categories.

While these defined domains help to articulate and compartmentalize functional requirements within program management, the analysis examines relationships both within and across these specific domains. In a broad sense, intra-domain evaluations measure the degree of coupling and integration among constructs that serve similar functions, whereas inter-domain evaluations characterize the degree of alignment between the distinct functional domains. This evaluation approach enables a structured assessment of acquisition practices in terms of their level of integration and level of independence amongst paired constructs.

Regression analysis focused on several statistical measures that help to assess both the magnitude and reliability of the relationship constructs. The first were three forms of regression coefficients. The raw mean regression coefficient listed in each linear regression chart represents the expected change in the dependent variable (vertical axis) for a one-unit increase in the independent variable (horizontal axis) on the original ordinal (0-3) scale. This preserves interpretability and allows direct comparisons with the scoring framework used in earlier parts of this analysis. It is referred to as raw mean β in the analysis.

The second and primary form of regression coefficient (β) used was ordinary least squares (OLS). While the raw mean coefficient provides an intuitive approximation of the direction and strength of relationships based on the previous ordinal scoring, it does not account for variance and covariance within the data structure. The OLS regression coefficient is used as the primary measure of association and represents the best-fit linear relationship across all variables and observations. This allows for statistical inference using standard error, confidence intervals, and p-values. It is referenced as OLS β in the analysis.

The third is a standardized beta coefficient. These were derived from z-scored variables. This was done to facilitate the comparison of relative effect sizes across all models. It is referenced as STD β in the analysis. Together, these three measures most accurately capture descriptive approximation (Raw Mean β), statistically estimated relationships (OLS β), and normalized comparative strength measures (STD β).

Additionally, R^2 as a measure of how much variance in the dependent variable (Y/ vertical axis) is explained by independent variable (X / horizontal axis). Higher R^2 values indicate a stronger explanatory relationship.

Statistical reliability is evaluated in modeling using both p-values and confidence intervals. The p-value tests whether the observed relationship is distinguishable from zero, with p-values below 0.05 indicating statistically significant evidence of a relationship between the two tested variables. Confidence intervals provide a range of values for the true regression coefficient. This research elected to report confidence intervals at the 95% level, meaning it is 95% certain that the true regression coefficient lies between the lower and upper interval bands. Intervals that do not include zero or negative numbers indicate statistically significant relationships. Intervals that include zero or negative values suggest insufficient evidence to conclude that a meaningful relationship exists.

Collectively, these measures provide a complementary view of each relationship, where β reflects the size of the effect, R^2 reflects the explanatory power, and p-values and confidence intervals assess statistical reliability. All regression results are interpreted using this combined analysis framework.

K. EXPLORATION OF RELATIONSHIPS

The first set of relationships analyzed was the concept within the execution-integration domain. This consisted of three unique relationship pairings. The first was centered on the extent to which open systems architectures are associated with iterative development.

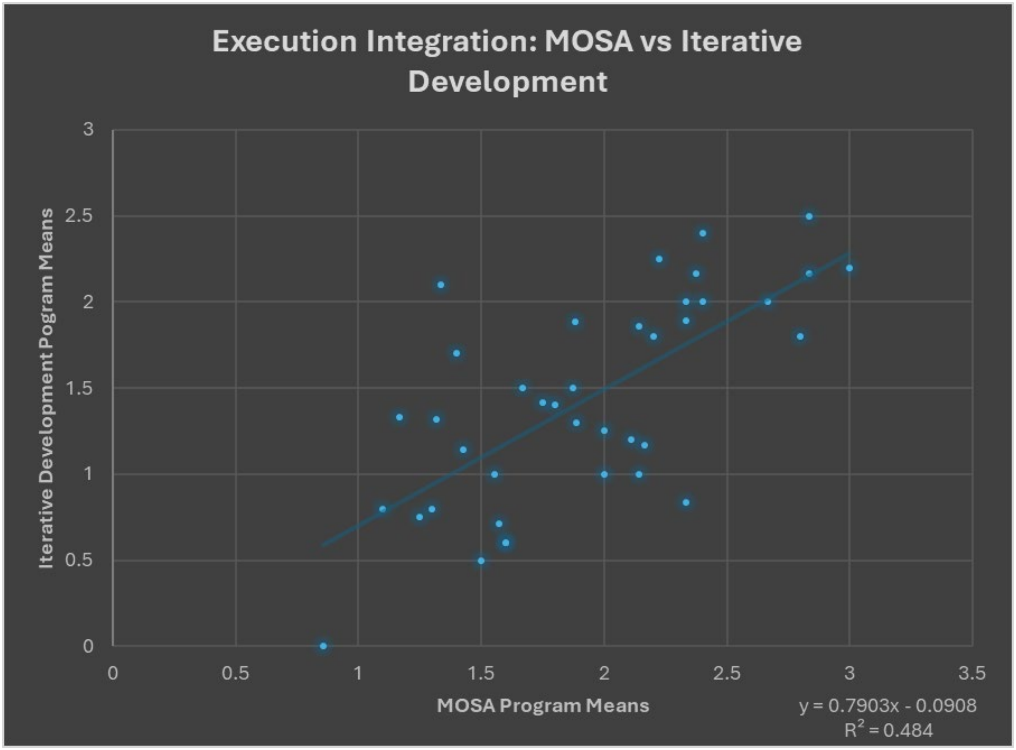

Figure 13. MOSA and iterative Development Linear Regression

The first relationship pair examined was MOSA's relation to iterative development (Figure 15). The relational hypothesis was that modular open architectures would be more likely to be present in programs with an incremental delivery structure, as open architectures enable flexible integration and ease of module replacement. This, in turn, would enable an incremental approach. This relationship showed a strong positive slope, with a raw mean regression coefficient of .79, indicating that a 1-point increase in MOSA is associated with a .79-point increase in iterative development. An R^2 of 0.484, which indicates that 48.4% of the variance in iterative development is explained by MOSA. The strong regression coefficient and high R^2 suggest that iterative development methodologies and open architectures are structurally aligned within the data set and may function as a coupled, mutually reinforcing execution strategy across many programs. Additionally, the confidence interval for this pairing was above zero and within a relatively tight range, with a spread of 0.421 between the upper and lower limits. This, along with a p-value below 0.001, indicates statistical significance.

Figure 14. M&S and Iterative Development Linear Regression

The second relationship pair was Modeling and Simulation's association with Iterative Development (Figure 14). This pairing was of interest, as programmatic use of modeling and simulation might be more common in an agile or iterative environment. Modeling and simulation would, in theory, reduce risk and accelerate iterative development cycles. The relationship has a moderate-to-strong positive slope, with a raw regression coefficient of 0.496 and moderate explanatory power ($R^2 = 0.223$). The confidence interval for this pairing was above zero, with a spread of 0.558; the reported p-value was 0.0024, and both measures were statistically significant. This suggests that the use of modeling and simulation techniques such as digital twins, analytical models, and virtual environments is at least somewhat associated with iterative development methods.

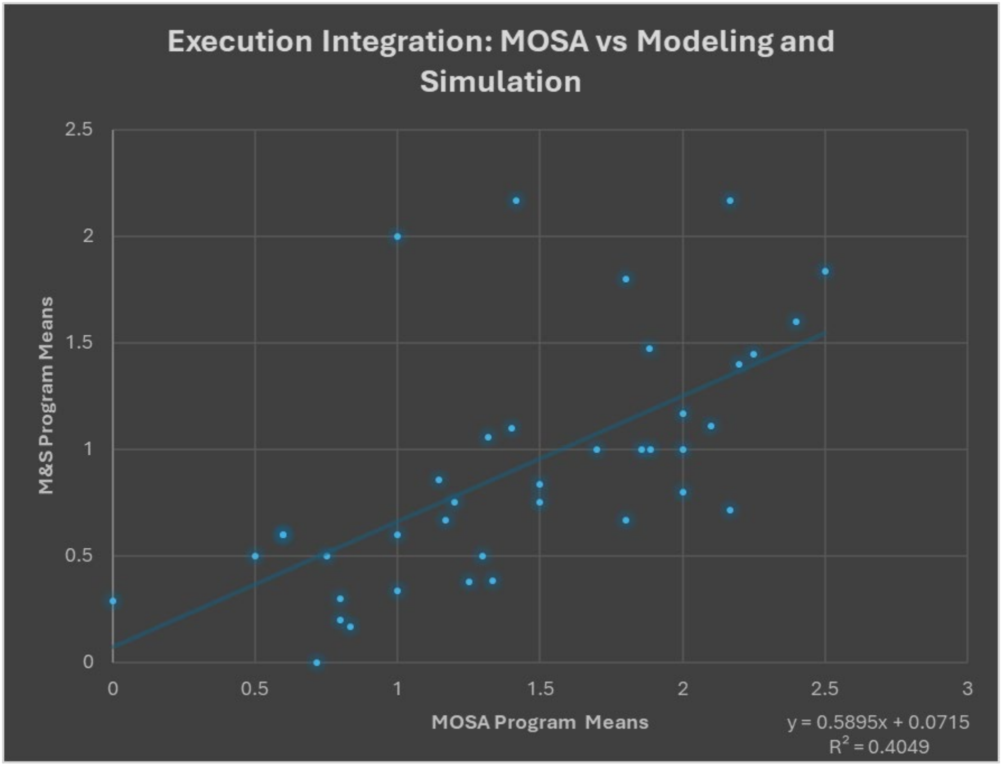

Figure 15. MOSA and M&S Linear Regression

The third execution integration testing theme was MOSA's association with Modeling and Simulation (Figure 15). The theory of potential relationships was grounded in the supposition that open system architectures would be more likely to employ modeling and simulation, as programs with a high MOSA emphasis would place greater weight on system decomposition and interface testing. Modeling and Simulation would perhaps be an important tool in component-level development and as an efficient, cost-effective way to test interface arrangements. The relationship demonstrated a strong positive slope, a coefficient of 0.590, and a moderately strong R^2 of 0.405. The confidence interval was above zero and exhibited a 0.477 spread between upper and lower limits. The p-value was below 0.001. This suggests that open systems architectures are strongly associated with the use of modeling and simulation techniques.

Figure 16. TRL and Business Case Linear Regression

Within the second analysis area of risk management and decision-making structures within MTARP program documentation (Figure 16). The relationship between TRL and the business case was examined. The hypothesized relationship logic is that technological maturity informs cost estimates and risk trade-offs; both concepts are closely related to what would be articulated in a business case analysis. This relationship should be positive, as higher technical maturity should inform a stronger business case in terms of cost and performance. There was a moderately positive slope with a coefficient of 0.443 and a moderate-to-weak R^2 of 0.150. The confidence interval was above zero but had a wide spread of 0.702, indicating greater uncertainty about the true coefficient (OLS β). The p-value of 0.0147 indicates a statistically significant relationship. This suggests that TRL could contribute to the strength of the business case, but it is not the dominant driver of successful MTARP programs.

Figure 17. TRL and Iterative Development Linear Regression

The next risk management and decision-making themed relationship was between technological readiness level and iterative development (Figure 17). The central assumption of this relationship is that technical maturity is closely associated with program risk, and iterative development reflects a desire to manage uncertainty and assumes an evolutionary path for integration of technologies. Therefore, programs with lower TRL scores might involve more iterative development. It is important to note that this hypothesized relationship would have a negative slope, unlike the previous examples, all of which had expected positive slopes. The results, however, showed a weak-to-moderate positive slope, suggesting that the relationship between higher iterative development and lower TRL was not present in the dataset. Further evidence to support a lack of relationship is found in the confidence interval, whose lower bound was a negative value, and a p-value of 0.21, suggesting no significant relationship. Finally, an R^2 of 0.042 indicates that this relationship has almost no explanatory utility.

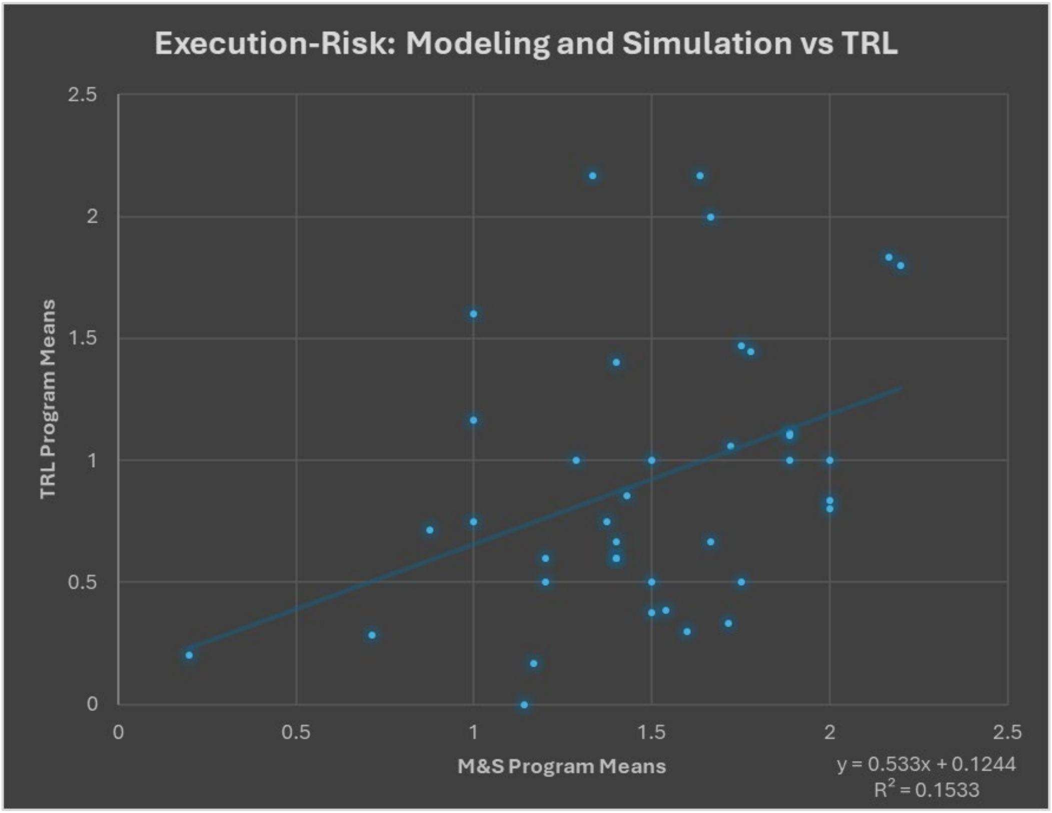

Figure 18. M&S and TRL Linear Regression

The final regression model involving risk management and decision-making examines the impact of modeling and simulation on technological readiness level (Figure 18). The implied relationship is that higher levels of modeling and simulation would correspond to higher technological readiness levels within the data set. Said another way, programs that implement more modeling and simulation tend to have higher technological readiness levels. The resulting linear regression indicates a moderate-to-strong positive slope, with a coefficient of 0.533: for every 1-unit increase in M&S, there would be a 0.533-unit increase in TRL using ordinal scoring scale. The R^2 of 0.153, however, is a moderate-to-weak explanatory quotient, explaining only 15.3% of the variance, suggesting that other significant factors are not accounted for in this model. In terms of statistical significance, the confidence interval was above zero, with a spread of 0.450 and a p-value of 0.014, both suggesting a significant relationship between the two constructs beyond mere chance.

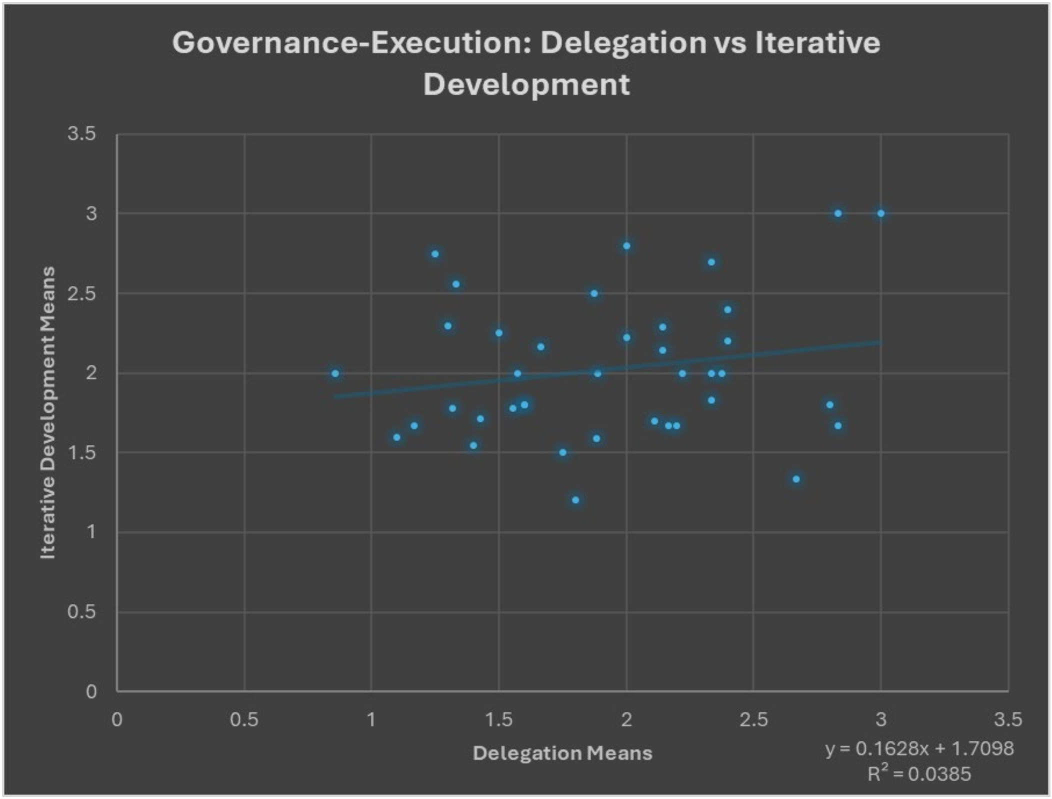

Figure 19. Delegation of Authority and Iterative Development Linear Regression

The third thematic exploration examined the association between programmatic governance and execution factors, measuring relationships with iterative development and open systems architecture (Figure 19). The first relationship pair was the delegation-of-authority relationship to iterative development. The hypothesized relationship was that iterative development often requires flexibility and course adjustments along the development path; that personnel at lower levels of authority would, in theory, be best positioned to make these adjustments; and that delegation of authority would be positively associated with iterative development. The results of the regression, however, indicate a weak positive slope coefficient of 0.163 and a very weak R^2 of 0.039. The confidence interval included negative values on the lower end of the interval, and had a p-value of 0.231, suggesting that no meaningful relationship exists between these two constructs.

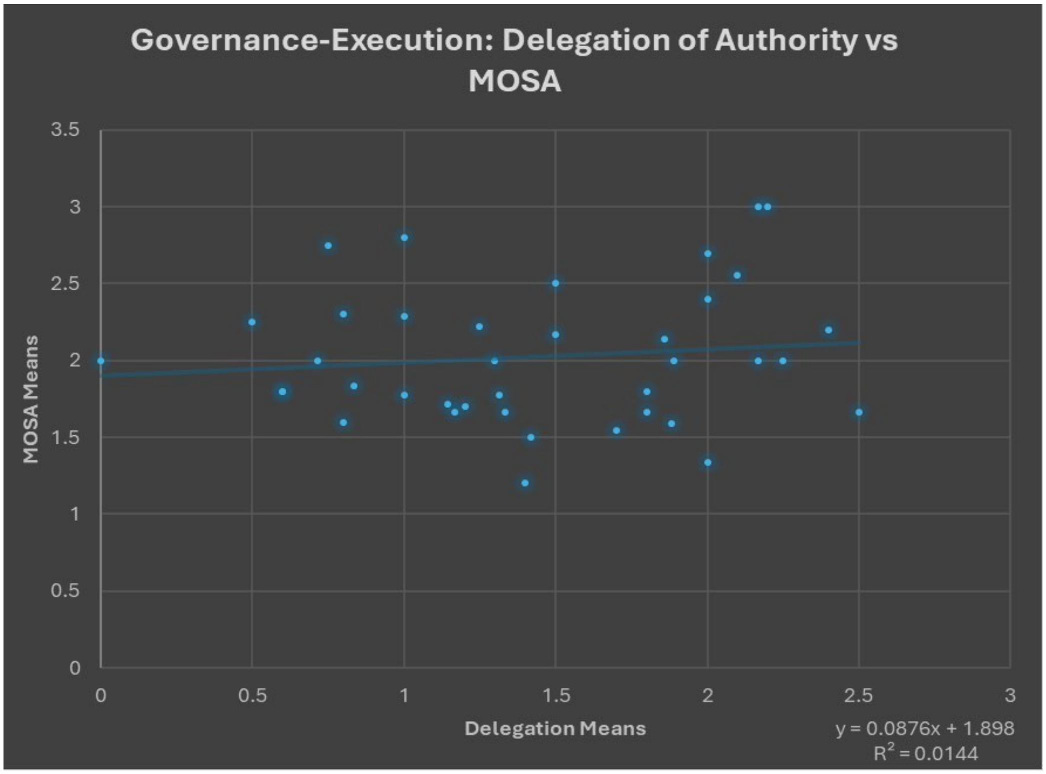

Figure 20. Delegation of Authority and MOSA Linear Regression

The final pairing within the governance realm and final simple linear regression exploration is the relationship of authority delegation to modular open systems architecture (Figure 20). The hypothesized explanation is that delegated authority might be more likely to be adopted in programs that institute MOSA. This is because MOSA adoption suggests that there is latitude in system architecture, interface, and modularity choices that will need to be made during the program lifecycle. Delegated authority would enable more informed decisions regarding system architecture. The regression results yielded a weak positive slope coefficient of 0.087 and an extremely weak R^2 of 0.014, indicating no relationship or explanatory power between delegated authority and MOSA. Further, the confidence interval included negative values and had a p-value of 0.467, implying no relationship exists.

L. CROSS-MODEL STATISTICAL INTERPRETATION

Earlier descriptive estimates of relationships (referred to as "raw mean beta") were used to approximate directional effects between variables using the original ordinal scoring method from structured prompts. Cross-model interpretation will use OLS coefficients, as this β estimator accounts for the data's variance (spread within a single variable) and covariance (how two variables move in relation to each other), yielding statistically valid estimates for inference. Additionally, the report standardized z-scored beta to facilitate uniform comparison of effect sizing across all variable pairs.

Table 4. Regression Analysis of MTA Best Practice Alignment

Regression Relationship (X → Y)	OLS β	STD β	SE	R^2	95% CI [Lower, Upper]	P-value
MOSA → Iterative Development	0.6125	0.6957	0.10396	0.4840	[.402, .823]	p < 0.001
M&S → Iterative Development	0.4482	0.4717	0.13775	0.2225	[.169, .727]	0.0024
MOSA →M&S	0.5895	0.6363	0.1174	0.4049	[.351, .828]	p < 0.001
TRL → Business Case	0.4431	0.3875	0.1733	0.1502	[.092, .794]	0.0147
TRL→ Iterative Development	0.2657	0.2054	0.2081	0.0422	[-.156, .687]	0.2097
Modeling and Simulation → TRL	0.2877	0.3916	0.1111	0.1533	[.063, .513]	0.0137
Delegation → Iterative Development	0.2366	0.1963	0.1944	0.0385	[-.157, .630]	0.2311
Delegation → MOSA	0.0876	0.1199	0.2235	0.0144	[-.289, .617]	0.4671

The regression results indicate that execution-integration related constructs exhibit the strongest relationships within the dataset. MOSA demonstrated a strong positive association with both iterative development (OLS β = 0.613, $R^2 = 0.484$, $p < 0.001$) and modeling and simulation (OLS β = 0.590, $R^2 = 0.405$, $p < 0.001$), suggesting a tight integration between open-systems architecture and iterative development methodologies. Similarly, modeling and simulation show a still strong but less so than MOSA, and statistically significant relationship with iterative development (OLS β = 0.448, $R^2 = 0.223$, $p = 0.002$), reinforcing the concept of digital engineering architectures in support of iterative execution. These findings collectively indicate that architectural flexibility and analytical capability are strongly aligned with iterative development approaches.

In the middle tier of strength and significance was the relationship between modeling and simulation and technological readiness levels. The relationship was positive and statistically significant (OLS β = 0.288, $R^2 = 0.153$, $p = 0.014$), with the confidence interval remaining above zero. This indicates that an increased use of modeling and simulation is associated with higher levels of technical maturity. However, R^2 is relatively modest, suggesting that modeling and simulation are not a dominant driver of technological readiness. Instead, TRL is likely influenced by a broad set of factors, with modeling and simulation simply one of many mechanisms that support TRL determination.

Governance-related variables exhibited weak and statistically insignificant relationships across all pairings. Delegation of authority shows negligible association with both MOSA (OLS $\beta = 0.088$, $R^2 = 0.014$, $p = 0.467$) and iterative development (OLS $\beta = 0.237$, $R^2 = 0.039$, $p = 0.231$), suggesting that programmatic governance structures may not be primary drivers of architectural decisions or program delivery approaches within the examined programs. Risk management variable TRL presented mixed results, demonstrating a moderate relationship with business case strength (OLS $\beta = 0.443$, $R^2 = 0.150$, $p = 0.015$), but a weak and non-significant association with iterative development (OLS $\beta = 0.266$, $R^2 = 0.042$, $p = 0.210$). These findings suggest that while technical maturity contributes to program justification, it does not strongly determine execution approach, and that architectural and analytical factors play a more central role in shaping program behavior.

Across all relationship models, standardized beta (STD β) was closely aligned with OLS, and z-scoring the raw means did not materially change the interpretation or strength of the relationships. This is likely due to the readily comparable ordinal scaling from which all data were derived and the inherent limitations on variability, which constrain the limit of scaling differences on the total data set. While standardized beta does not change substantive interpretation of results, it does provide a normalized basis for comparison and reinforces the conclusions drawn from OLS beta interpretations.

M. MULTIPLE LINEAR REGRESSION

The exploration of pairwise relationships among the best-practice constructs, using a three-program management functional-area framework, provides an initial assessment of how execution, risk/decision, and governance variables interact within the program dataset. Some of our simple linear regression pairings did demonstrate statistically significant associations. However, these results do not indicate whether the relationships are truly independent, reflect overlap among related construct variables, or control for other potential sources of variance. Multiple linear regression allows for greater isolation of relationships among variables and helps determine which observed associations persist after controlling for other related constructs, which variables contribute independent explanatory power, and which do not. For this analysis, there were four selected sets of multivariate relationships.

The first multivariate relationship explored linked the previously identified significance between iterative development and MOSA and sought to determine if the addition of the execution variable of M&S and the risk management variable of TRL would more fully explain the drivers of iterative development, or if it was primarily related to open systems architecture and MOSA principles. This multivariate relationship was statistically significant ($F = 11.09$, $p < 0.001$), but there was no meaningful increase in the R^2 measure of explanatory power between the simple linear regression between MOSA and iterative development ($R^2 = 0.484$) and the multivariate statistic ($R^2 = 0.487$). Individual coefficients (β) and p values, for the three independent variables, support the limited change in R^2 where the MOSA coefficient remains as the strongest and statistically significant predictor of iterative development ($\beta = 0.592$, $p < 0.001$), while both M&S ($\beta = 0.060$, $p = 0.701$) and TRL $\beta = -0.061$, $p = 0.724$) were not statistically significant and include confidence intervals that include zero.

Table 5. Multivariate Regression 1

$\varepsilon + (M\&S)\beta_3 + (MOSA)\beta_2 + (TRL)\beta_1 + \beta_0 =$ Iterative Development			
$R^2 = .443$	F = 11.08	P-Value < 0.001	SE = 0.399
	β	P-Value	95% CI [Upper, Lower]
M&S	0.060	0.701	[-0.252, 0.371]
MOSA	0.592	< 0.001	[.309, .874]
TRL	-0.061	0.724	[-0.409, 0.287]

These findings reveal that, while M&S demonstrated a moderate pairwise relationship in the simple linear regression (OLS $\beta = 0.448$, $R^2 = 0.223$, $p < 0.024$), its explanatory power does not hold when controlling for MOSA. Whereas TRL's relationship to iterative development did not have a meaningful relationship (OLS $\beta = 0.266$, $R^2 = 0.042$, $p < 0.210$) this did remain consistent within the multivariate regression, despite an assumed hypothetical connection between the use of analytical simulation and technological maturity and their influence on iterative development strategies. The lack of explanatory value for these variables suggests that TRL does not have a meaningful relationship to iterative development, and the M&S relationship to iterative development is not a potentially independent explanatory factor as first surmised in simple linear regression. The result also reinforces the significance of the relationship between MOSA and iterative development as having real explanatory power.

Table 6. Multivariate Regression 2

$\varepsilon + (\text{Iterative})\beta_2 + (\text{MOSA})\beta_1 + \beta_0 = \text{M\&S}$			
$R^2 = .407$	F = 12.33	P-Value < 0.001	SE = 0.446
	β	P-Value	95% CI [Upper, Lower]
Iterative	0.059	0.755	[-0.322, 0.440]
MOSA	0.553	0.002	[0.217, 0.889]

The next multivariate regression was completed to further examine the relationship between M&S and its statistically significant relationship to iterative development (OLS $\beta = 0.448$, $R^2 = 0.223$, $p = 0.024$), and MOSA's statistically significant relationship to M&S (OLS $\beta = 0.590$, $R^2 = 0.405$, $p < 0.001$). The multivariate regression used iterative development and MOSA as the independent variables, and M&S as the dependent variable, with the purpose of determining if M&S had a direct relationship to iterative development or if there is significant covariation between MOSA and M&S. The resulting multivariate relationship was statistically significant ($F = 12.33$, $p < 0.001$) however there was little change in between the multivariate R^2 (0.407) and the simple linear regression R^2 (0.405) of MOSA and M&S. The individual coefficients (β) indicate that MOSA continues to be a strong and statistically significant predictor of M&S ($\beta = 0.553$, $p = 0.002$) and that iterative development does not possess a statistically significant relationship ($\beta = 0.059$, $p = 0.755$). This suggests that M&S's relationship to iterative development does not have strong explanatory power beyond what MOSA accounts for in relation to iterative development.

Table 7. Multivariate Regression 3

$\varepsilon + (\text{MOSA})\beta_2 + (\text{M\&S})\beta_1 + \beta_0 = \text{TRL}$			
$R^2 = .167$	F = 3.61	P-Value = 0.037	SE = 0.388
	β	P-Value	95% CI [Upper, Lower]
MOSA	0.103	0.447	[-0.169, 0.375]
M&S	0.217	0.143	[-0.077, 0.5106]

The third multivariate regression model was built to test the soundness of the three-area program management framework conceptualized from best-practice constructs, and the decision to separate TRL as a risk/decision function rather than an execution function, even though it could reasonably be argued that it could be either. This analysis measured the association between the TRL risk management construct and the execution constructs of MOSA and M&S. The hypothesis was that there would be no statistically significant relationship, because TRL is not an embedded execution construct but rather a distinct component of program management (risk/decision construct). This hypothesis would stand in contrast to what is already known about the pairwise relationships among the execution constructs (MOSA, M&S, and iterative development), all of which have statistically significant relationships, as evidenced by the first three simple linear regressions conducted in the prior section. The multivariate regression was statistically significant at the aggregate level ($F = 3.61$, $p = 0.037$), but the explanatory power of the R^2 was relatively weak (0.167). The individual independent variables' coefficients (β) reveal that neither MOSA ($\beta = 0.103$, $p = 0.447$) nor M&S ($\beta = 0.217$, $p = 0.143$) are statistically significant, and both confidence intervals include zero. Even though the earlier pairwise regression between M&S and TRL suggested a moderate and statistically significant relationship (OLS $\beta = 0.288$, $R^2 = 0.153$, $p = 0.013$) this relationship no longer holds once controlling for MOSA, and the results suggest that TRL is not strongly associated with either MOSA or M&S, and stands in contrast to the strong and consistent relationships between MOSA, M&S, and iterative development in prior pairwise comparisons. TRL appears to operate independently from these constructs and gives a level of validation as a separate conceptual framework within the risk/decision realm of program management.

Table 8. Multivariate Regression 4

$\varepsilon + (MOSA)\beta_3 + (TRL)\beta_2 + (Iterative)\beta_1 + \beta_0 = $ Business Case			
$R^2 = .193$	F = 2.79	P-Value = 0.055	SE = 0.443
	β	P-Value	95% CI [Upper, Lower]
Iterative	0.151	0.424	[-0.229, 0.531]
TRL	0.381	0.047	[0.006, 0.756]
MOSA	0.044	0.8	[-0.304, 0.392]

The final multivariate analysis asked which other variables might explain the business case variable. The lone simple linear regression, with the business case as the dependent variable, was TRL's relationship to business case. This relationship was statistically significant and moderately explanatory (OLS $\beta = 0.443$, $R^2 = 0.150$, $p = 0.015$). The multivariate analysis will add execution-oriented constructs of MOSA and iterative development to determine whether any additional explanatory power or covariance can be uncovered. The results of the multivariate model explained marginally more ($R^2 = 0.193$)

than did TRL's relationship to business case and approached the cusp of statistical significance ($F = 2.79$, $p = 0.055$), suggesting limited total explanatory power of this multivariate model. The individual coefficients further show that TRL's relationship to business case remains statistically significant ($\beta = 0.381$, $p = 0.047$) as it was in the pairwise regression. Iterative development ($\beta = 0.151$, $p = 0.424$) and MOSA ($\beta = 0.044$, $p = 0.800$) were not statistically significant and contained confidence intervals below zero. These results give additional legitimacy to the conceptual framework where risk/decision constructs of TRL and business case were assumed to be structurally embedded, although the relationship is modest, the lack of relationship with execution-based constructs does lend additional credence to the conceptual soundness of the risk/decision construct as being a distinct entity within the wider best practice program management constructs.

N. CONSTRAINTS AND LIMITATIONS

This research is subject to several important limitations that affect the interpretation and generalizability of the findings. These limitations fall into three main categories: data constraints, methodological constraints, and analytical constraints.

The completeness and consistency of the data within the Acquisition Information Repository (AIR) were significant constraints. Program documentation varied significantly, with many programs having minimal artifacts. This was controlled by eliminating 45 successful and 15 unsuccessful programs from the analysis due to minimal artifact data. This is not necessarily a failure of AIR, but, based on observed patterns in the dataset, it is a feature of non-major MTA programs more broadly. There are minimal documentation requirements for these types of programs, and all the successful programs eliminated fell into the non-major system acquisition category. Additional data constraints come from the use of placeholder documents within various programs for classified materials. These placeholder documents were systematically removed from analysis because they did not add value or insight into programmatic practices. The final data constraint came from corrupted files within AIR that could not be extracted, viewed, or analyzed. These were also removed from analysis when discovered.

Methodological constraints were impacted by dependence on GenAI.mil's Gemini LLM to sort the large volume of program data. The use of a rubric to constrain interpretation significantly mitigates risks associated with hallucination by requiring verbatim text extraction and skewed inference by requiring structured scoring. It cannot fully mitigate the model's interpretation of acquisition program language. There are real potential constraints on semantic variability. For example, concepts that an acquisition-knowledgeable human reader would consider simple phrasing differences may receive different scores from the model depending on its interpretation. Finally, even structured prompts with specified rubrics are subject to interpretation drift. The LLM will approximate the specified boundaries and will still have to make judgment calls about what is potentially included and what is not. This, however, does not differ significantly from the potential pitfalls of human-led inference and interpretation of rule-based constructs.

The second methodological constraint that the examined program documents can only demonstrate written intent to operationalize best practices within the program elements; they are not evidence of execution of those best practices within the development effort. That level of clarity would likely require structured interviews with program participants and is outside the scope of this analysis.

The first analytical constraint is the use of ordinal scoring to rate program documents. This does allow for structured comparison across all qualitative constructs. Ordinal scoring reduces nuanced qualitative information to simplified categories and may distort measurement validity. The second analytical constraint centered around sample size, and specifically the sample size of unsuccessful programs. With only six unsuccessful programs meeting minimum documentation requirements for inclusion, statistical comparisons involving unsuccessful outcomes were not completed. This severely limits statistical comparisons between the two groups.

Finally, the regression analysis was exploratory in nature. It assessed pairwise relationships (via simple regression) between best-practice constructs but did not control for additional variables that might influence the documented programmatic intent. As such, these results identify associations between constructs, and not causal relationships, and should be framed within the broader context of acquisition program complexity. These findings do provide directional and relational insights into acquisition best practice intent but should not be interpreted as causal explanations of program intent.

V. SUMMARY, CONCLUSIONS, RECOMMENDATIONS

A. SUMMARY

This capstone research examined the extent to which DoD MTA rapid prototyping programs align with acquisition best practice constructs and how that alignment relates to program outcomes. The analysis answered a primary research question and four secondary questions:

1. Primary Research Question

- To what extent do DoD MTA rapid prototyping programs demonstrate documented alignment with selected acquisition best practice constructs, and how does that alignment vary across program status categories, Services, and best practice domains?

The findings indicate that DoD MTA rapid prototyping programs demonstrated varying levels of documented alignment with selected acquisition best practices constructs across AIR/DAVE program artifacts. Alignment differed across program status categories, military Services, and construct domains, programs associated with more favorable outcomes generally demonstrated stronger documented alignment in areas such as technology maturity and business-case-related practices.

2. Secondary Research Questions

- How consistently are selected acquisition best practice constructs documented across MTA rapid prototyping program artifacts in the AIR/DAVE dataset?

The analysis found inconsistent documentation of selected best practice constructs across AIR/DAVE program artifacts. Certain constructs, including iterative development and delegation-related practices, appeared more consistently than others, while some constructs demonstrated substantial variation in both frequency and depth of documentation.

- How does documented best practice alignment differ across AIR/DAVE program status categories, including programs identified as successful, transitioned, suspended, terminated, disposed, or otherwise unsuccessful?

Programs associated with successful or transitioned outcomes generally demonstrate stronger documented alignment across several best practice constructs when compared to programs categorized as suspended, terminated, disposed, or otherwise unsuccessful. However, alignment patterns varied by construct and were not uniformly predictive of program outcomes.

- How does documented best practice alignment vary across military Services within MTA rapid prototyping programs?

The findings identified observable variation in documented best practice alignment across the military Services. These differences suggest that Service-affiliated MTA rapid prototyping programs may emphasize acquisition practices differently within program documentation.

- What exploratory relationships exist among documented best practice constructs within MTA rapid prototyping documentation?

Exploratory quantitative analysis identified several moderate relationships among selected best practice constructs, suggesting that certain acquisition practices may function interdependently within MTA rapid prototyping efforts. While the analysis did not establish causation, the findings indicate potential associations among related acquisition practice domains.

The findings reveal variation in best practice alignment across MTA rapid prototyping programs and demonstrate that more favorable outcomes are associated with consistent and comprehensive alignment across multiple best practices. These results provide a foundation for understanding how practices may influence program performance with the MTA pathway.

B. CONCLUSIONS

This capstone research found that the selected best practices do not function as purely independent contributors to successful programs. Instead, the results indicated a differentiated and graduated structure among the examined construct best-practice variables. For instance, MOSA demonstrates consistently strong relationships with the other two execution constructs, M&S and iterative development. Additionally, risk and decision management constructs demonstrated a strong relationship with TRL and business case, being statistically significant and possessing a moderately explanatory R^2. Within the final construct domain of governance, as expressed through the delegation of authority variable, no statistically significant relationships were found when paired with the dependent variables of iterative development and MOSA. Where relationship patterns really began to break down were in cross-domain construct pairings (i.e., execution constructs paired with risk & decision management and/or governance). All of the examined cross-domain pairings, apart from M&S's relationship to TRL, were not statistically significant.

Among the multivariate regressions, MOSA's dominance of the execution construct domain becomes more apparent. In examining iterative development and MOSA's relationship with M&S, it became clear that MOSA accounts for the majority of this relationship, with a β of 0.553 and a statistically significant p-value, whereas iterative development had a β of only 0.059 and no statistically significant p-value. This suggests that the bivariate relationships of M&S and other execution variables are largely captured by MOSA, and that MOSA retains significant independent explanatory power and accounts for a meaningful portion of variance within the execution construct domain.

The inclusion of multiple variables did not produce more meaningful increases in overall model fit between the bivariate and multivariate regressions. For instance, the relationship between M&S and TRL yielded an R^2 of 0.153 and a p-value of 0.0137. When MOSA was added as an additional independent variable, the R^2 increased marginally to 0.167, and the p-value became less significant at 0.037. This was repeated between the bivariate relationship of TRL and business case (β 0.443, R^2 0.150, p-value 0.015), and the multivariate relationship of the independent variables MOSA, TRL, and iterative development and their relationship to business case (R^2 0.193, p-value 0.055) and the β for TRL at 0.381, indicating that TRL remained the dominant explanatory variable within the multivariate relationship.

Regarding the relationships between successful and unsuccessful MTA rapid prototyping programs, technology readiness at program entry appears to be an influential factor that differentiates successful from unsuccessful programs, although the risk of sampling error due to the low sample size of unsuccessful programs makes inferential conclusions imprudent. Though successful programs demonstrate a 57% higher TRL (Mean = 1.566) than unsuccessful programs (Mean = 0.997). While this finding is descriptive rather than inferential, it aligns with existing GAO assessments and reinforces that the compressed timelines inherent in MTA are generally more compatible with programs that begin with sufficiently mature technologies.

While the results of this research do not support causal claims for or against the effectiveness of MTA rapid prototyping, instead, they do indicate that the examined best-practice constructs exhibit domain-bounded patterns of association rather than an integrated cross-domain system of relationships. A potential implication of these findings is that evaluating MTA rapid prototyping programs using aggregated or cross-domain evaluation criteria might obscure how these various constructs operate in practice. This suggests that future oversight and assessment frameworks could benefit from placing greater emphasis on domain-specific alignments rather than cross-domain evaluation frameworks.

C. RECOMMENDATIONS

In many ways, the implementation of the Middle Tier of Acquisition as part of the 2016 NDAA served as a harbinger of DoD acquisition realignment, prioritizing rapid capability development and delivery. Accordingly, past MTA efforts contain significant experiential insight, process information, and generated data that have the potential to inform the DoD's shift away from centralized JROC-oriented requirements controls towards leaner Service-led portfolio-based acquisition. This transformation is underway across all of the Services, and has led the Department of the Navy to stand up five Portfolio Acquisition Executive (PAE) organizations, covering major naval domains such as undersea, maritime, and Marine Corps portfolios, as well as more institutional domains like industrial operations, and strategic deterrence with the strategic systems portfolio (Department of the Navy, 2026). Likewise, the Army has aligned its capability development around six new PAEs: maneuver ground, maneuver air, fires, sustainment and ammo, layered protection, and command and control portfolios (Marino, 2026). Similar realignments are underway within the Air and Space Forces as well.

However, documentation requirements for MTA programs to date have been sparse, and with good reason. They were intended to be as lean as possible, so as not to inhibit program execution, add administrative burden, or slow progress towards delivering capabilities to meet additional documentation requirements. The hindsight risk with this approach is the probability that the Services have been conducting rapid acquisition experiments without a robust data-collection methodology capable of determining which characteristics, attributes, and decisions meaningfully impact cost, schedule, and performance of MTA programmatic efforts. Future collection and methodical analysis are critical to developing a stronger evidence base for MTA and would be equally critical to implementing evidence-based refinements and reforms that meaningfully shorten development cycles and deliver capabilities sooner.

For future MTA program efforts, an empirical evaluation of a wide variety of program management aspects should be conducted. Longitudinal MTA program development data could be an area where Operational Research/Systems Analysis (ORSA) professionals or similarly qualified personnel could evaluate program performance, transition outcomes, cost and schedule behavior, and downstream operational effects. This analysis should not only include whether a given MTARP transitioned, but also why the MTA pathway was initially selected, what the intended end state was at program initiation, and the sponsor's perception of value through the lens of viable operational capability delivery prior to transition. The variability between these factors amongst program efforts would likely provide a more meaningful basis for distinguishing among MTARP outcomes. Those that truly accelerated capability delivery, those that appropriately used the pathway to reduce technical risk and requirements uncertainty before transitions, and those that ultimately just shifted risks into other acquisition pathways.

For program efforts that have reached completion, a systematic review of successfully transitioned MTARP programs would perhaps be a good starting point. These reviews might draw on program information beyond what is available in AIR and might even include retrospective interviews to clarify intent and execution decisions made during an MTARP program lifecycle. This data might be collected across various program management domains, similar to how the research is presented in this capstone, and it should include aspects beyond this capstone, such as cost estimates versus realized costs, sponsor/warfighter perceptions of value, and post-transition performance and system integration outcomes.

Without systematic data collection and analysis, the DoD will be unable to adequately determine whether MTA and the related but much more recent Service-level portfolio-based reforms are actually accelerating capability delivery or whether they are simply shifting requirements development, technical maturation, system integration risk, and production capability risk reduction onto different system development vehicles and organizational structures.

D. RECOMMENDED AREAS FOR FUTURE RESEARCH

The findings of this study highlight both the value and the limitations of analyzing MTA programs through documented alignment with acquisition best practice constructs, indicating several important areas for future research. While this research provides insight into how constructs are reflected in program documentation, a more comprehensive understanding of MTA effectiveness will require expanded datasets, additional constructs, and deeper examination of how these practices are implemented in execution.

One important area for future research is the expansion of this analysis across a larger and more longitudinal dataset. Applying a similar structured methodology to a broader population of MTA programs over time would enable a more robust statistical analysis and improve the ability to identify consistent patterns, trends, and potential indicators of program success or risk. Incorporating additional data points would enhance the ability to assess MTA effectiveness and support the development of a scalable, metrics-based evaluation framework. The additional data points could include execution outcomes, schedule performance, transition success, and tension in TRL/MRL levels.

Building on the constructs examined, future research should also consider the inclusion of additional readiness and structural variables that may influence program outcomes. One such variable is Manufacturing Readiness Level (MRL), which represents a critical dimension of a program's ability to transition from development to production. Manufacturing Readiness is becoming increasingly prescient as we continue to embark upon the transformation of the Warfighting Acquisition System and meet aggressive production delivery schedules and rapid delivery of incremental improvements (Office of the Secretary of War [OSW], 2025, p. 1). The SECWAR Transformation Memorandum dated 7 November 2025 explicitly calls for the establishment of a Wartime Production Unit as an office within the Office of the Undersecretary of War for Acquisition and Sustainment (OUSW [A&S]) to improve coordination between the military departments and defense industry partners in support of a faster delivery of capabilities (OSW, 2025, p. 1-2). The criticality of the subfactors of MRL, namely industrial base readiness, process capability and control, facilities and tooling maturity, and supply chain resilience, should all be important assessed factors in future MTA endeavors, especially given the emphasis on accelerating production of capabilities to the Warfighter (Office of the Secretary of Defense Manufacturing Technology Program [OSD ManTech], 2025, p. 5-13). The Manufacturing Readiness Level Deskbook provides a suggested entry point of MRL 4 for Rapid Prototyping Programs and MRL 8 for Rapid Fielding (OSD ManTech, 2025, p. 27-29). These suggested levels could be used as a starting point for the development of structured LLM prompts and subsequent data set exploration.

Another potential construct is the role of intellectual property (IP) management, particularly in implementing MOSA. While MOSA is often framed as a technical implementation strategy focusing on interoperability and adaptability, it is more accurately understood as an integrated business and technical strategy, in which contract decisions specifically related to IP governance and data rights, directly shape architectural outcomes and, by extension, the effectiveness of MOSA implementation (Rendon, 2006, p. 10). Future research could examine how different IP strategies influence system integration, competition, and long-term adaptability, and whether variations in IP management contribute to differences in how modularity is implemented across programs. This would be of particular interest as MOSA appears to be the structural driver of the execution constructs (MOSA, M&S, and iterative development) within the previous simple linear regressions. Adding an exploration of Intellectual Property strategies in future research with similarly structured designs might help further distinguish programs that claim MOSA modularity but lack the IP strategy to achieve true open architectures from those that have higher-level alignment with MOSA through IP conditioning. It would be interesting to see this distinction made and the separate MOSA (strong IP vs weak IP) constructs tested again against the execution constructs of M&S and iterative development.

In addition to expanding constructs, future research should explore organizational and human factors that may influence program outcomes. While this study focused on documented acquisition practices, factors such as leadership experience, program manager decision-making, governance structures, and workforce capability may significantly shape how programs interpret and apply the flexibility inherent in the MTA pathway. Understanding how these elements interact with acquisition practices would provide a more holistic perspective on MTA execution.

These areas of future research would support a more comprehensive and system-level understanding of MTA performance. By expanding the scope of analysis to include additional variables, broader datasets, and deeper examination of execution, future studies can build on the foundation established in this research and contribute to the development of informed, data-driven acquisition decision-making.

APPENDIX A. BUSINESS CASE LLM PROMPT

SYSTEM INSTRUCTION:

You are an AI research assistant performing structured document analysis. Your task is to analyze the provided document to identify and score all evidence related to the program's business case, focusing on its dynamism and risk management aspects. You must adhere strictly to the provided rubric and output format.

USER PROMPT:

Analyze the entire document provided and identify every section that contains evidence of a sound business case. For each distinct instance you identify, you must evaluate it using the following rubric and output format.

Rubric:

- Score 0: No mention of business case, affordability analysis, life-cycle cost, cost-benefit analysis, Analysis of Alternatives (AoA), cost targets, or trade-offs between cost, schedule, and performance.
- Score 1 (Superficial): A program establishes top-level cost goals or affordability caps but provides no detail on the analysis that justifies them or the strategy to achieve them.
- Score 2 (Intent + Partial Mechanism): The program explicitly discusses the trade-space between cost, schedule, and performance. It may reference the completion of an Analysis of Alternatives (AoA) or a formal cost estimate as the basis for its plan.
- Score 3 (Operationalized / Strong): The business case is presented as a dynamic tool. There is clear evidence of a strategy to manage risk and make decisions, such as: defining specific criteria for "off-ramps" or re-evaluation points, linking technical maturation to funding decisions, or a plan to use early prototyping to validate the business case before committing to full-scale development.

Instructions:

1. Analyze the Full Document: Scan the entire document for all qualifying evidence.
2. Assign a Single Overall Score: Based on all the evidence found, assign one single, overall score (0–3) for the entire document.
3. Write an overall Justification: Provide a brief summary justification that explains why the overall score was given.
4. Detail Each Instance: For each piece of evidence, create a separate entry containing the direct quote, page number, and a specific justification tied

to the rubric.

5. Compile into a Single Output: The final response must be a single, cohesive block of text that follows the mandatory format below.
6. Do Not Paraphrase: The evidence must be a direct quote.
7. No Evidence: If no qualifying evidence is found, output only the "Overall" block with a score of 0 and the justification "No qualifying Business Case evidence found in the document."

Mandatory Output Format:

Overall Score: <0|1|2|3>

Overall Justification:

Instance 1

All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

Instance 2

All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

(Repeat the Instance block for each additional piece of evidence found)

APPENDIX B. DELEGATION OF AUTHORITY LLM PROMPT

SYSTEM INSTRUCTION:

You are an AI research assistant performing structured document analysis. Your task is to analyze the provided document to identify and score all mentions of the delegation of decision authority. You must adhere strictly to the provided rubric and output format, focusing on delegation below the Service Acquisition Executive (SAE).

USER PROMPT:

Analyze the entire document provided and identify every section that contains evidence of delegated decision authority. For each distinct instance you identify, you must evaluate it using the following rubric and output format.

Rubric:

- Score 0: No mention of the Milestone Decision Authority (MDA), decision authority, delegated authorities, or the chain of command for program decisions.
- Score 1 (Superficial / Standard Structure): Mentions the Service Acquisition Executive (SAE) as the Milestone Decision Authority (MDA) but provides no evidence of further delegation to the PEO or PM level.
- Score 2 (Intent + Partial Mechanism): Explicitly states intent to empower the PM/PEO or streamline decisions. May mention the PEO having some decision-making role without citing specific, formal delegation of cost, schedule, or performance trade-space authority.
- Score 3 (Operationalized / Strong): Provides clear and concrete evidence of authority being delegated below the SAE. This includes citing a specific Acquisition Decision Memorandum (ADM) that delegates authority, or explicitly stating the PM/PEO has been granted authority for requirements trades, funding realignment, or other key programmatic decisions.

Instructions:

1. Analyze the Full Document: Scan the entire document for all qualifying evidence.
2. Assign a Single Overall Score: Based on all the evidence found, assign one single, overall score (0–3) for the entire document.
3. Write an Overall Justification: Provide a brief summary justification that explains why the overall score was given.
4. Detail Each Instance: For each piece of evidence, create a separate entry containing the direct quote, page number, and a specific justification tied to the rubric.

5. Compile into a Single Output: The final response must be a single, cohesive block of text that follows the mandatory format below.
6. Do Not Paraphrase: The evidence must be a direct quote.
7. No Evidence: If no qualifying evidence is found, output only the "Overall" block with a score of 0 and the justification "No qualifying Delegation of Authority evidence found in the document."

Mandatory Output Format:

Overall Score: <0|1|2|3>

Overall Justification:

\---

Instance 1

All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

\---

Instance 2

All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

\---

(Repeat the Instance block for each additional piece of evidence found)

APPENDIX C. ITERATIVE APPROACH LLM PROMPT

SYSTEM INSTRUCTION:

You are an AI research assistant performing structured document analysis. Your task is to analyze the provided document to identify and score all mentions of Iterative Development and its related concepts. You must adhere strictly to the provided rubric and output format.

USER PROMPT:

Analyze the entire document provided and identify every section that contains evidence of an Iterative Development approach. For each distinct instance you identify, you must evaluate it using the following rubric and output format.

Rubric:

- Score 0: No mention of iterative development, spiral development, increments, agile methods, Minimum Viable Product (MVP), DevSecOps, soldier touchpoints, or a phased capability delivery approach.
- Score 1 (Superficial): A generic reference to "iterative development," "agile," or "spirals" with no specific details on schedules, feedback mechanisms, or delivery increments.
- Score 2 (Intent + Partial Mechanism): Explicit intent to use an iterative approach plus at least one of the following: a high-level schedule of increments/spirals, a defined plan for user feedback (e.g., Soldier Touch Points), or a strategy for delivering partial capabilities over time.
- Score 3 (Operationalized / Strong): A clear and detailed iterative plan with concrete enabling mechanisms such as: a defined schedule of capability drops or software releases, established procedures for continuous user feedback and integration, linkage of development cycles to test events, or a formal Agile methodology (e.g., Scrum, Kanban) being specified.

Instructions:

1. Analyze the Full Document: Scan the entire document for all qualifying evidence. Do not include simple mentions of Rapid Prototyping and Rapid Fielding on their own as they are universally applicable to larger acquisition constructs of the documents examined. If they are mentioned along with iterative principles, then they may be included.
2. Assign a Single Overall Score: Based on all the evidence found, assign one single, overall score (0–3) for the entire document.
3. Write an Overall Justification: Provide a brief summary justification that explains why the overall score was given.

4. Detail Each Instance: For each piece of evidence, create a separate entry containing the direct quote, page number, and a specific justification tied to the rubric.
5. Compile into a Single Output: The final response must be a single, cohesive block of text that follows the mandatory format below.
6. Do Not Paraphrase: The evidence must be a direct quote.
7. No Evidence: If no qualifying evidence is found, output only the "Overall" block with a score of 0 and the justification "No qualifying Iterative Development evidence found in the document."

Mandatory Output Format:

Overall Score: <0|1|2|3>

Overall Justification:

Instance 1

All Available Evidence: "<Direct quote(s) from document>"

Page:
Justification:

Instance 2

All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

(Repeat the Instance block for each additional piece of evidence found)

APPENDIX D. MODELING & SIMULATION LLM PROMPT

SYSTEM INSTRUCTION:

You are an AI research assistant performing structured document analysis. Your task is to analyze the provided document to identify and score all mentions of Modeling and Simulation (M&S) and its related concepts. You must adhere strictly to the provided rubric and output format.

USER PROMPT:

Analyze the entire document provided and identify every section that contains evidence of a Modeling and Simulation approach. For each distinct instance you identify, you must evaluate it using the following rubric and output format.

Rubric:

- Score 0: No mention of modeling, simulation, M&S, digital twin, digital engineering, digital thread, virtual environment, virtual prototype, hardware-in-the-loop (HWIL), or software-in-the-loop (SWIL).
- Score 1 (Superficial): A generic reference to using "modeling and simulation" or "digital engineering" with no specific application, tool, or purpose mentioned.
- Score 2 (Intent + Partial Mechanism): Explicit intent to use M&S plus a description of its application for a specific purpose, such as: requirements analysis, design trade studies, supplementing test events, or creating a virtual prototype.
- Score 3 (Operationalized / Strong): A clear and detailed M&S plan with concrete enabling mechanisms such as: identifying specific M&S tools or environments, linking M&S to formal design reviews (e.g., SRR, PDR, CDR), use of a digital twin for performance prediction, or integrating M&S (e.g., HWIL/SWIL) as part of the formal test and evaluation strategy.

Instructions:

1. Analyze the Full Document: Scan the entire document for all qualifying evidence. Do not include simple mentions of Rapid Prototyping and Rapid Fielding on their own as they are universally applicable to larger acquisition constructs of the documents examined. If they are mentioned along with iterative principles, then they may be included.
2. Assign a Single Overall Score: Based on all the evidence found, assign one single, overall score (0–3) for the entire document.
3. Write an Overall Justification: Provide a brief summary justification that

explains why the overall score was given.

4. Detail Each Instance: For each piece of evidence, create a separate entry containing the direct quote, page number, and a specific justification tied to the rubric.
5. Compile into a Single Output: The final response must be a single, cohesive block of text that follows the mandatory format below.
6. Do Not Paraphrase: The evidence must be a direct quote.
7. No Evidence: If no qualifying evidence is found, output only the "Overall" block with a score of 0 and the justification "No qualifying Modeling and Simulation evidence found in the document."

Mandatory Output Format:

Overall Score: <0|1|2|3>

Overall Justification:

Instance 1

All Available Evidence: "<Direct quote(s) from document>"

Page:
Justification:

Instance 2

All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

(Repeat the Instance block for each additional piece of evidence found)

APPENDIX E. MOSA LLM PROMPT

SYSTEM INSTRUCTION:

You are an AI research assistant performing structured document analysis. Your task is to analyze the provided document to identify and score all mentions of Modular Open Systems Architecture (MOSA) and its related concepts. You must adhere strictly to the provided rubric and output format.

USER PROMPT:

Analyze the entire document provided and identify every section that contains evidence of a MOSA approach. For each distinct instance you identify, you must evaluate it using the following rubric and output format.

Rubric:

- Score 0: No mention of MOSA, open systems, open architecture, modular design, open interfaces, standards-based architecture, or competition-enabling architectural mechanisms.
- Score 1 (Superficial): A generic reference to modularity or open architecture with no specific mechanisms, standards, interfaces, or implementation details.
- Score 2 (Intent + Partial Mechanism): Explicit MOSA or open systems intent plus at least one of the following: defined open interfaces, standards-based architecture, modular decomposition strategy, interoperability emphasis, or upgradeability language.
- Score 3 (Operationalized / Strong): Clear MOSA commitment with concrete enabling mechanisms such as: interface control documents (ICDs), defined interface standards profiles, government purpose rights (GPR) or data rights enabling competition, a specific technical data package (TDP) strategy, or a vendor interoperability strategy.

Instructions:

1. Analyze the Full Document: Scan the entire document for all qualifying evidence.
2. Assign a Single Overall Score: Based on all the evidence found, assign one single, overall score (0–3) for the entire document.
3. Write an Overall Justification: Provide a brief summary justification that explains why the overall score was given.
4. Detail Each Instance: For each piece of evidence, create a separate entry containing the direct quote, page number, and a specific justification tied to the rubric.

5. Compile into a Single Output: The final response must be a single, cohesive block of text that follows the mandatory format below.
6. Do Not Paraphrase: The evidence must be a direct quote.
7. No Evidence: If no qualifying evidence is found, output only the "Overall" block with a score of 0 and the justification "No qualifying MOSA evidence found in the document."

Mandatory Output Format:

Overall Score: <0|1|2|3>

Overall Justification:

Instance 1

All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

Instance 2
All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

(Repeat the Instance block for each additional piece of evidence found)

APPENDIX F. TECHNOLOGY READINESS LEVEL LLM PROMPT

SYSTEM INSTRUCTION:

You are an AI research assistant performing structured document analysis. Your task is to analyze how the provided document reports on and manages technology readiness. You must adhere strictly to the provided rubric and output format, focusing on the distinction between system-level and component-level technology readiness.

USER PROMPT:

Analyze the entire document provided and identify every section that contains evidence of technology readiness assessment and maturation. For each distinct instance you identify, you must evaluate it using the following rubric and output format.

Rubric:

- Score 0: No mention of Technology Readiness Level (TRL), critical technologies, technology maturation, or technology risk.
- Score 1 (Superficial / Opaque): Reports a single Technology Readiness Level (TRL) at the overall system or program level. Fails to identify the individual critical technologies that make up the system.
- Score 2 (Intent + Partial Mechanism): Identifies and lists the program's specific critical technologies, but either reports their TRLs collectively or does not provide a TRL for each one. May mention a "Technology Maturation Plan" without detailing the specific activities or milestones.
- Score 3 (Operationalized / Strong): Provides clear and transparent evidence of rigorous technology management. This includes a breakdown of multiple critical technologies, an individual TRL assigned to each one, and a corresponding plan (e.g., a Technology Maturation Plan) that details the specific efforts to advance the readiness of any immature technologies.

Instructions:

1. Analyze the Full Document: Scan the entire document for all qualifying evidence.
2. Assign a Single Overall Score: Based on all the evidence found, assign one single, overall score (0–3) for the entire document.
3. Write an Overall Justification: Provide a brief summary justification that explains why the overall score was given.
4. Detail Each Instance: For each piece of evidence, create a separate entry containing the direct quote, page number, and a specific justification tied to the rubric.

5. Compile into a Single Output: The final response must be a single, cohesive block of text that follows the mandatory format below.
6. Do Not Paraphrase: The evidence must be a direct quote.
7. No Evidence: If no qualifying evidence is found, output only the "Overall" block with a score of 0 and the justification "No qualifying Technology Maturation evidence found in the document."

Mandatory Output Format:

Overall Score: <0|1|2|3>

Overall Justification:

Instance 1

All Available Evidence: "<Direct quote(s) from document>"

Page:

Justification:

Instance 2

All Available Evidence: "<Direct quote(s) from document>"
Page:

Justification:

(Repeat the Instance block for each additional piece of evidence found)

LIST OF REFERENCES

Anton, P. S., Tannehill, B., McKeon, J., Goirigolzarri, B., Holliday, M. A., Lorell, M. A., & Younossi, O. (2020). *Strategies for acquisition agility: Approaches for speeding delivery of defense capabilities* (RR-4193). RAND Corporation. https://www.rand.org/pubs/research_reports/RR4193.html

Bartels, E. M., Drezner, J. A., & Predd, J. B. (2020). *Building a broader evidence base for defense acquisition policymaking* (RRA-202-1). RAND Corporation. https://www.rand.org/pubs/research_reports/RRA202-1.html

Blumenson, M. (1993). *Breakout and Pursuit*. Center of Military History, United States Army.

Bub, K. (2023). *An analysis of rapid prototyping with the DoD* [Master's capstone project, Naval Postgraduate School]. NPS Archive: Calhoun https://dair.nps.edu/bitstream/123456789/4971/1/NPS-AM-23-199.pdf

Candreva, P. J. (2023). National defense budgeting and financial management: Policy and practice (2nd ed.). Information Age Publishing.

Chairman of the Joint Chiefs of Staff. (2026, January 15). *Manual for the Joint Requirements Oversight Council and the Joint Force Requirements Process* (CJCSM 5123.01). https://www.jcs.mil/Portals/36/Documents/Library/Manuals/ CJCSM%205123.01.pdf

Cohen, M. A., Eliashberg, J., & Ho, T.-H. (1996). New product development: The performance and time-to-market tradeoff. *Management Science*, *42*(2), 173–186.

Defense Acquisition University. (2023). *ACQ 0150 Middle tier of acquisition: Course student guide*. Defense Acquisition University. https://www.dau.edu/courses/acq-0150

Defense Acquisition University. (2026). *ACQ 0010: Welcome to Defense Acquisition.* [Presentation]. Defense Acquisition University. https://www.dau.edu/courses/acq-0010

DeLuca, P., Predd, J. B., Nixon, M., Blickstein, I., Button, R. W., Kallimani, J. G., & Tierney, S. (2013). *Assessing Aegis program transition to an open-architecture model*. RAND Corporation. https://www.rand.org/pubs/research_reports/ RR161.html

Department of Defense Office of Inspector General. (2021). *Audit of Department of Defense middle tier of acquisition rapid prototyping and rapid fielding programs* (DODIG-2021-131). https://media.defense.gov/2021/Sep/30/2002864712/-1/-1/1/ DODIG-2021-131.PDF

Department of the Air Force, Office of the Assistant Secretary of the Air Force for Acquisition, Technology & Logistics. (2021, May 7). *DoDI 5000.80 / DAFI 63-146: Operation of the middle tier of acquisition (MTA)*. [Memorandum] Department of the Air Force. https://static.e-publishing.af.mil/production/1/ saf_aq/publication/dodi5000.80_dafi63-146/dodi5000.80_dafi63-146.pdf

Department of the Army, Office of the Assistant Secretary of the Army (Acquisition, Logistics and Technology). (2018, September 25). *Office of the Assistant Secretary of the Army (Acquisition, Logistics and Technology) middle-tier acquisition policy* [Memorandum]. Department of the Army. https://aida.mitre.org/wp-content/uploads/2018/10/MTA-ASAALT-Policy-25-Sep-18-3.pdf

Department of the Army, Office of the Assistant Secretary of the Army (Acquisition, Logistics and Technology). (2020, March 20). *Office of the Assistant Secretary of the Army (Acquisition, Logistics and Technology) middle tier of acquisition (MTA) policy* [Memorandum]. Department of the Army. https://aida.mitre.org/ wp-content/uploads/2020/04/ASAALT-MTA-Policy.pdf

Department of the Navy, Office of the Secretary of the Navy. (2022, April 8).

Department of the Navy implementation and operation of the Defense Acquisition System (Change transmittal 1) (SECNAVINST 5000.2G). [Memorandum] Department of the Navy. https://www.secnav.navy.mil/doni/Directives/ 05000%20General%20Management%20Security%20and%20Safety%20Services/ 05000%20General%20Admin%20and%20Management%20Support/5000.2G.pdf

Drezner, J. A., McKernan, M., Sollinger, J. M., & Newberry, S. (2020). *Issues with access to acquisition information in the Department of Defense* (RR-3130). RAND Corporation. https://www.rand.org/pubs/research_reports/RR3130.html

Etemadi, A., & Kamp, J. (2022a). Middle tier acquisitions and innovation. *Proceedings of the Nineteenth Annual Acquisition Research Symposium: Acquisition Research—Creating Synergy for Informed Change* (SYM AM 22 067). Naval Postgraduate School, Acquisition Research Program.

Etemadi, A., & Kamp, J. (2022b). Estimating middle tier acquisition schedule risk.

Proceedings of the Nineteenth Annual Acquisition Research Symposium: Acquisition Research—Creating Synergy for Informed Change (SYM AM 22 093). Naval Postgraduate School, Acquisition Research Program.

Fried, T. (2025, November 13). *The impact of drones on the battlefield: Lessons of the Russia-Ukraine war from a French perspective*. Hudson Institute. https://www.hudson.org/missile-defense/impact-drones-battlefield-lessons-russian-ukraine-war-french-perspective-tsiporah-fried

Government Accountability Office. (2016). Technology readiness assessment guide: Best practices for evaluating the readiness of technology for use in acquisition programs and projects (GAO-16-410G). https://www.gao.gov/assets/gao-16-410g.pdf

U.S. Government Accountability Office. (2020a). Cost estimating and assessment guide: Best practices for developing and managing program costs (GAO-20-195G).

Government Accountability Office. (2020b). *Weapon systems annual assessment: Defense acquisitions* (GAO-20-439). https://www.gao.gov/assets/gao-20-439.pdf

Government Accountability Office. (2022). *Weapon systems annual assessment: Challenges to fielding capabilities faster persist* (GAO-22-105230). https://www.gao.gov/assets/gao-22-105230.pdf

Government Accountability Office. (2023a). Middle-tier defense acquisitions: Rapid prototyping and fielding requires changes to oversight and development approaches (GAO-23-105008). https://www.gao.gov/assets/820/817501.pdf

Government Accountability Office. (2023b). Weapon systems annual assessment: Programs are not consistently implementing practices that can help accelerate acquisitions (GAO-23-106059). https://www.gao.gov/assets/gao-23-106059.pdf

Government Accountability Office. (2024). *Weapon systems annual assessment: DoD is not yet well-positioned to field systems with speed* (GAO-24-106831). https://www.gao.gov/assets/gao-24-106831.pdf

Government Accountability Office. (2025a). Adaptive acquisition framework: DOD should take steps to institutionalize iterative development practices across pathways (GAO-25-107003). https://www.gao.gov/assets/gao-25-107003.pdf

Government Accountability Office. (2025b). Weapon systems annual assessment: DOD leaders should ensure that newer programs are structured for speed and innovation (GAO-25-107569). https://www.gao.gov/assets/gao-25-107569.pdf

Information Technology Acquisition Advisory Council & Foundation for Defense of Democracies. (2025, August 8). A roadmap for sustainable IT acquisition reform 2.0: A digital transformational guide for the 47th president and executive branch leadership (EO 14265 public comment submission).

International Council on Systems Engineering. (2015). INCOSE systems engineering handbook: A guide for system life cycle processes and activities (4th ed.). John Wiley & Sons.

LaCamera, J. J., Jr. (2019, May 1). *Rapid acquisition: The challenge to accelerate* [PowerPoint slides]. National Defense Industrial Association. https://www.ndia.org/-/media/sites/ndia/divisions/ipmd/2019-04-meeting/207-middle-tier-of-acquisition-lacamera-190508215620.ashx

Leonard, R. S., & Wallace, A. (2014). Air Force major defense acquisition program cost growth is driven by three space programs and the F-35A: Fiscal year 2013 President's budget selected acquisition reports (RR-477-AF). RAND Corporation. https://www.rand.org/pubs/research_reports/RR477.html

Liu, N. F., K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2023).

Lost in the middle: How language models use long context. arXiv. https://arxiv.org/abs/2307.03172

Marino, C. (2026, March 11). *The Army's 2025 acquisition reforms revolutionize processes to expedite cutting-edge capabilities*. U.S. Army. https://www.army.mil/article/290080/the_armys_2025_acquisition_reforms_revolutionize_processes_to_expedite_cutting_edge_capabilities

McElheran, K. (2013). *Delegation in multi-establishment firms: Evidence from IT purchasing* (Working Paper No. 11-101). Harvard Business School.

McGarry, B. W. (2022). DOD planning, programming, budgeting, and execution (PPBE): Overview and selected issues for Congress (CRS Report No. R47178). Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/ R47178

McGarry, B. W. (2022). *DOD planning, programming, budgeting, and execution (PPBE): Overview and selected issues for Congress* (CRS Report No. R47178). Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/ R47178

McKernan, M., Drezner, J. A., Arena, M. V., Wong, J. P., Shokh, Y., Lewis, A., Moore, N. Y., Mele, J. D., & Newberry, S. (2022). *Using metrics to understand the performance of the adaptive acquisition framework* (RRA-1349-1). RAND Corporation. https://www.rand.org/pubs/research_reports/RRA1349-1.html

National Defense Authorization Act for Fiscal Year 2016, Section 804: Middle Tier of Acquisition for Rapid Prototyping and Rapid Fielding, Pub. L. 114-82 (2015). https://www.congress.gov/114/plaws/publ92/PLAW-114publ92.pdf

National Defense Authorization Act for Fiscal Year 2016, Pub. L. No. 114-92, 129 Stat. 726 (2015, November 25). https://www.congress.gov/114/plaws/publ92/PLAW-114publ92.pdf

National Defense Authorization Act for Fiscal Year 2018, Pub. L. No. 115-91, 131 Stat. 1283 (2017, December 12). https://www.congress.gov/115/plaws/publ91/PLAW-115publ91.pdf

National Defense Industrial Association. (2023). *Modular open systems approach: Implementation challenges and opportunities*. https://www.ndia.org/-/media/sites/ndia/divisions/systems-engineering/reports/2023/mosa-implementation-challenges-and-opportunities.ashx

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2019, December 30). *DoD Instruction 5000.80: Operation of the middle tier of acquisition (MTA) (Change 1, November 25, 2024)*. U.S. Department of Defense. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500080p.PDF

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020a, January 23). *DoD Instruction 5000.02: Operation of the adaptive acquisition framework (Change 1, June 8, 2022)*. U.S. Department of Defense. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500002p.PDF

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020b, September 9). *DoD Directive 5000.01: The defense acquisition system*. U.S. Department of Defense. https://www.esd.whs.mil/Portals/54/Documents/DD/ issuances/dodd/500001p.pdf

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020c, October 2). *DoD Instruction 5000.87: Operation of the software acquisition pathway*. U.S. Department of Defense. https://www.esd.whs.mil/Portals/54/ Documents/DD/issuances/dodi/500087p.PDF

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020d, August 6). *DoD Instruction 5000.85: Major capability acquisition*. U.S. Department of Defense. https://www.esd.whs.mil/Portals/54/Documents/DD/ issuances/dodi/500085p.pdf

Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics. (2017, February 2). DoD Instruction 5000.02: Operation of the Defense Acquisition System (Incorporating Change 2) [Enclosure 14: Cybersecurity in the Defense Acquisition System]. U.S. Department of Defense. https://dodprocurementtoolbox.com/uploads/ Do_DI_5000_02_Enclosure_14_Cybersecurity_in_the_Defense_Acquisition_System_6931c75bd5.pdf

Office of the Secretary of Defense Manufacturing Technology Program. (2025, May 1). Manufacturing readiness level (MRL) deskbook (Version 2025). https://www.dodmrl.com

Office of the Under Secretary of Defense for Research and Engineering, Prototypes and Experiments. (2022, October). *Department of Defense prototyping guidebook* (Version 3.1). U.S. Department of Defense.

Office of the Under Secretary of Defense for Research and Engineering. (2025). *Technology readiness assessment guidebook* (February 2025). Department of Defense.

Office of the Secretary of War. (2025, November 7). Transforming the defense acquisition system into the warfighting acquisition system to accelerate fielding of urgently needed capabilities to our warriors [Memorandum].

Perdue, W. M., & Teissonniere-Almodovar, T. I. (2024). *Adaptive acquisition framework: Effectiveness of the middle tier of acquisition pathway* (Master's capstone project, Naval Postgraduate School). NPS Archive: Calhoun https://dair.nps.edu/bitstream/123456789/5175/1/NPS-AM-24-188.pdf

Rendon, R. G. (2006). Using a modular open systems approach in defense acquisitions: Implications for the contracting process (NPS-AM-06-010). Naval Postgraduate School.

United States Special Operations Command. (2020). *USSOCOM Directive 71-4: Special Operations Forces Capabilities Integration and Development System (SOFCIDS)*. https://www.dau.edu/sites/default/files/Migrated/CopDocuments/Directive%20-%20USSOCOM%2C%2071-4%2C%20SOFCIDS.pdf

Webster, J. D. (2024). *The middle tier of acquisition: Speed vs. rigor*. Defense Acquisition University https://apps.dtic.mil/sti/html/trecms/AD1230720/ index.html

Acquisition Research Program
Department of Acquisition, Finance and Manpower
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net