SYM-AM-21-065

Excerpt from the Proceedings of the Eighteenth Annual Acquisition Research Symposium

Exploring Performance in Science and Technology Programs

May 11–13, 2021

Published: May 10, 2021

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

Acquisition Research Program
Graduate School of Defense Management
Naval Postgraduate School

The research presented in this report was supported by the Acquisition Research Program of the Graduate School of Defense Management at the Naval Postgraduate School.

Exploring Performance in Science and Technology Programs

Johnathan Ritschel—Air Force Institute of Technology

Scott Drylie—Air Force Institute of Technology

Clay Koschnick—Air Force Institute of Technology

Eric Plack—Air Force Cost Analysis Agency

Edward White—Air Force Institute of Technology

Abstract

Science and Technology (S&T) programs serve an important function in the defense acquisition process as the initial phase leading to discovery and development of warfighting technology. The results of these programs impact the larger Major Defense Acquisition Programs that integrate the technologies in subsequent phases of the lifecycle. Despite this important role, little prior research has examined the performance of S&T programs. Therefore, we investigate the impact of technological maturation as a critical success factor in S&T programs. The results suggest that S&T programs with mature technologies are more likely to experience above average cost growth and larger contract values while being less likely to experience schedule growth. Additionally, we find the partnership method between the government and contractor matters for both technological maturation and schedule growth. Lastly, the nature of the S&T program is important, with aerospace programs more likely to technologically mature than human systems programs.

An Analysis of Science and Technology Program Performance

Program management focuses on cost, schedule, and performance as the three key measures of success (Meredith & Mantel, 2003; Pinto & Slevin, 1998). A large body of literature identifies critical factors that lead to program success in both private industry (Nasir & Sahibuddin, 2011; Pinto & Slevin, 1987; Zwikael & Globerson, 2006) and the public sector (Rendon, 2012; Rodriguez-Segura et al., 2016; Tishler et al., 1996). Prior analyses of program performance in defense programs, however, have focused almost exclusively on larger, more mature programs that have reached the Engineering Manufacturing Development (EMD) phase of the lifecycle or beyond. An abundance of studies exploring cost growth or schedule growth can be found for these Major Defense Acquisition Programs (MDAPs; Bolten et al., 2008; Cancian, 2010; Smirnoff & Hicks, 2008). Missing from the literature is an exploration of smaller programs that feed basic science and technologies to subsequent acquisition programs or that develop new systems and technologies on a smaller scale. These are the Science and Technology (S&T) programs that are undertaken in defense research labs. This article seeks to bridge that gap through an exploratory analysis of program performance in Air Force S&T programs.

Importance of Science and Technology

The vision to implement science and technology as a centerpiece of our nation's airpower strategy has been around since 1945 (Duffner, 2000). General H. H. "Hap" Arnold, commanding general of the Army Air Forces, enlisted the aid of leading aeronautics scientist Dr. Theodore von Karman to lead the first of these efforts, recommending the creation of an agency devoted exclusively to aeronautical research and development (Gorn, 1988). Over time, that agency has evolved to what is known today as the Air Force Research Laboratories (AFRL; Duffner, 2000).

S&T's enduring importance is demonstrated in the 2019 publication of the Air Force Science and Technology Strategy for 2030. The 2030 S&T strategy aligns with the National

Defense Strategy to empower S&T programs to develop and deliver warfighting capabilities rapidly and effectively (U.S. Air Force, 2019). How does S&T fulfill this need? S&T functions as the initial phase of the acquisition process by which technologies are matured and, where appropriate, transitioned for acquisition by the Air Force (Office of the Chief Scientist of the U.S. Air Force, 2010). Continual advancement in these cutting-edge technologies is crucial, as the Air Force faces ever-changing threats and adversarial advancements in technology.

The Anatomy of Air Force Research Labs

The S&T data analyzed in this paper are from AFRL programs. A brief organizational description is provided for those unfamiliar with the laboratories. AFRL is headquartered at Wright-Patterson Air Force Base (AFB) in Ohio. It is comprised of nine technology directorates in the continental United States and four locations overseas in Hawaii, the United Kingdom, Chile, and Japan, as shown in Figure 1.

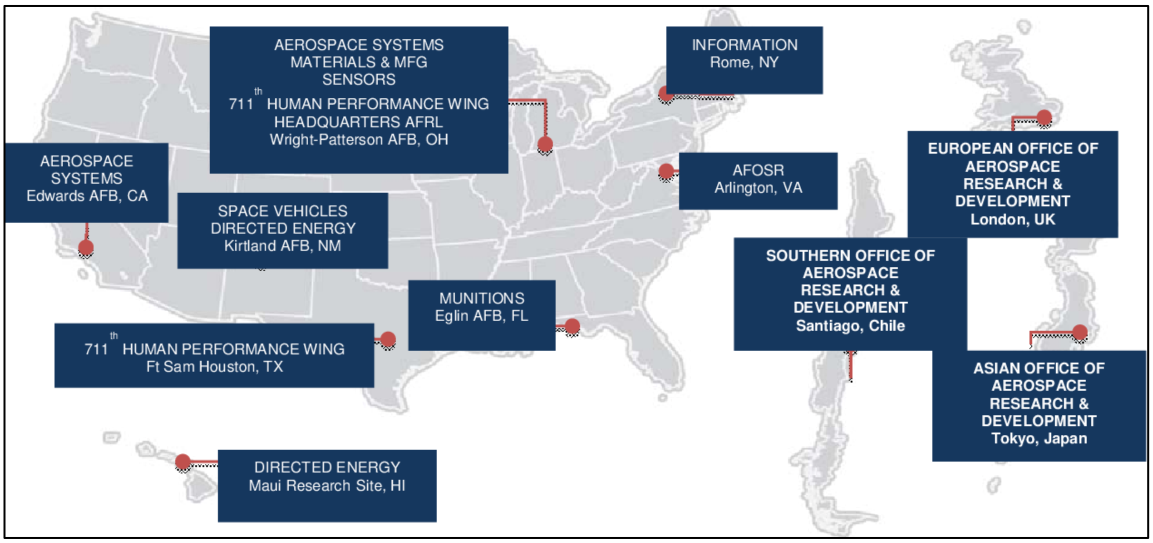

Figure 1: AFRL Locations and Major Offices

Each technology directorate focuses on the development and innovation of leading-edge technologies and is separated by technological capabilities. A list of AFRL's technology directorates, office symbols, and program descriptions are provided in Table 1. The analysis of individual technical directorates will be one of the ways this research segments the data.

Table 1: AFRL Technology Directorates

Technology Directorate	Symbol	Program Descriptions
Air Force Office of Scientific	AFOSR	Basic Research Manager for AFRL
711th Human Performance Wing	RH	Aerospace Medicine, Human Systems Integration
Directed Energy Directorate	RD	Laser, Electromagnetics, Electro-Optics
Information Directorate	RI	Information Fusion, Exploitation, Networking
Aerospace Systems Directorate	RQ	Aerodynamics, Flight Control, Engines, Propulsion
Space Vehicle Directorate	RV	Space-Based Surveillance, Capability Protection
Munitions Directorate	RW	Air-Launched Munitions
Materials & Manufacturing	RX	Aircraft, Spacecraft, Missiles, Rockets
Sensors Directorate	RY	Sensors for Reconnaissance, Surveillance

Measures of Success: The Role of Technology Readiness Levels

The General Accounting Office (GAO) has identified technology maturation as a critical success factor in product development (GAO, 1999). The Department of Defense's (DoD's) approach to incorporate this critical success factor has been to emphasize Technology Readiness Levels (TRLs) as a measure for selecting mature technologies for inclusion in a program (DoD, 2011). The TRL concept was developed by NASA (Sadin et al., 1989) and has subsequently been adopted by AFRL. A TRL is a tool to measure the technology maturity of a system or subsystem using a nine-level ordinal scale (DoD, 2011). Detailed TRL definitions and descriptions can be found in Appendix A.

It is believed that "programs that enter the Engineering and Manufacturing Development (EMD) phase of the Defense Acquisition System and have immature technologies will incur cost growth and schedule slippage" (DoD, 2009). In an effort to reduce the risk associated with entering the EMD phase of the acquisition lifecycle at Milestone B, DoD Instruction 5000.02 requires technologies to be demonstrated in a relevant environment (i.e., obtain a TRL of at least 6; DoD, 2011). AFRL, through the S&T programs it oversees, serves a key role in the creation and maturation of these technologies to reach those thresholds.

Despite TRLs being identified as a critical success factor, the literature is sparse with empirical examinations. The dearth of analysis is particularly acute for S&T type programs, but even MDAPs have relatively few studies examining TRLs. Dubos et al. (2008) analyzed the relationship between technology uncertainty and schedule slippage in the space industry. Their research resulted in the creation of TRL-schedule-risk curves that are intended to assist program managers to make informed decisions regarding the appropriate TRL to consider when confronted with schedule constraints. The research of Dubos et al. (2008) suggested a close relationship between technology uncertainty and schedule risk, where the more mature a technology is (the higher the TRL), the less potential schedule slippage.

Katz et al. (2015) specifically studied the relationship of TRLs to cost and schedule changes during the EMD phase. They found that weapon systems that achieved a TRL of 7 or greater at Milestone B had a lower probability of schedule slippage during the EMD phase than weapons systems that had a TRL of less than 7. While Katz et al. (2015) found evidence to suggest that technology maturity is related to schedule change, they did not find a relationship with cost changes.

Smoker and Smith (2007), however, found evidence that suggests costs vary exponentially across time as the system's technology progresses through each TRL. Similarly, Linick (2017) found that as the TRL increased throughout the development phase, the percentage of the development cost increased at an increasing rate. As shown by the literature, the extant TRL studies are primarily focused on programs once they reach the EMD stage. To the best of our knowledge, there are no studies that focus solely on S&T programs—a gap this paper is designed to fill.

Data

The data for this research was obtained from the AFRL cost and economics division. S&T programs typically fall below the dollar threshold for traditional standardized reporting such as Contract Performance Reports (CPRs). Instead, the S&T programs receive Funds and Man-Hour Expenditure Reports (FMERs). These FMERs provide the procuring activity visibility into the contractor's expenditures for labor, materials and parts, travel, subcontractors, and other charges. Like CPRs, these reports are required on a periodic basis from the contractor, usually monthly. Unlike CPRs, FMERs do not report standardized cost elements like the ones found in MIL-STD-881D. The initial AFRL dataset consisted of 165 S&T programs with contract start dates spanning from 2009 to 2017.

Research Summary Reports were also collected for these programs. These reports are generated at the start of the program (Initial), during the program (Periodic), and at the end of the program (Final). Research Summary Reports include general information such as the program title, lead technical directorate, and start/end dates. They also include DoD-required information such as performance type, joint capability area, Air Force technical capabilities, and TRLs. An example of a Research Summary Report can be found in Appendix B.

Of the 165 programs obtained from AFRL, 43 are included in the final dataset. Table 2 provides the exclusion criteria and associated number of programs remaining in the analysis.

Table 2. Dataset Exclusions

Category	Number Removed	Remaining Programs
Programs Obtained from AFRL		165
Missing Elements	64	101
Inadequate TD Sample Size	10	91
Less Than 92.5% Complete	48	43
Final Dataset for Analysis		43

As shown in Table 2, programs which had missing elements are excluded. These 64 programs had their costs reported on the FMER in unique ways to include cost burn rates, earned value management graphs, total costs in phases, or simply an overall total cost or labor hours spent. These reporting methods lack the specific elements needed in this analysis to compute percentages of total cost which are used to observe the program's behavior. Of the 101 remaining programs, 10 programs fall under four different technical directorates (RD, RI, RX, and RY). Each technical directorate represents unique programs with different characteristics which precluded aggregation above the technical directorate level. Therefore, the small sample size in these directorates would likely skew the analysis results, especially when observing how these programs behave at the technical directorate level. For these reasons, the programs are excluded from the analysis. Finally, programs with a completion percentage of less than 92.5% are excluded from the dataset. A program's completion percentage is computed using the total cost from the last available FMER to the program's contract value at that time. Previous research determined that a program with a completion percentage of 92.5% or greater accurately predicts the final cost of the program (Tracy & White, 2011). The final number of programs in the dataset is 43, which is sufficient to conduct a robust analysis.

Methods: Contingency Table Analysis

The dataset consists largely of qualitative variables. Therefore, the methodological approach employed is a two-way contingency table analysis. This type of analysis is used to summarize the relationship between two categorical variables based on the data observed. The contingency table analysis uses a 2 × 2 table to test for independence. For each test, the same type of hypothesis test will be implemented, as shown in Equation 1:

H_o: The two classifications are independent

H_a: The two classifications are dependent (1)

The chi-square distribution is the test statistic used for considering inferences about the category probabilities. If there is a failure to reject the null, the two variables are independent and are not statistically related to one another. If the null is rejected, then the variables are dependent, and a statistical relationship exists between them. The two-way contingency analysis examines the categorical variables (see Table 3) with subsequent discussion on the rationale behind variable selection and categorization.

When highly significant results are found, one of the benefits of a contingency table is that odds ratios and their associated confidence intervals can be produced. An odds ratio is a measure of association for a two-way contingency table. The ratio is the odds of an event occurring in one group to the odds of the same event occurring in another group. In other words, the odds ratio is the ratio of the probability of a property being present compared to the probability of it being absent. If the odds ratio is 1, the two events are independent.

Table 3. Categorical Variables Used in Contingency Table Analysis

Categorical Variables	
Technical Directorate	Cost Growth > 0%
Performance Type	Cost Growth > 33.7%
TRL Increase	Cost Growth > 44.1%
Last Known TRL ≥ 6	Cost Growth > 56.5%
Final TRL ≥ 6	Cost Growth > 60.5%
TRL 1 – 3	Cost Growth > 68%
TRL 4 – 5	Contract Value > $1M
TRL 6 – 7	Contract Value > $3M
TRL 8 – 9	
Schedule Growth > 0%	
Schedule Growth > 33%	
Schedule Growth > 63%	

Categorical variables for the Technical Directorate (TD), Performance Type, and TRLs are obtained from the Research Summary Reports. The TD variable denotes which AFRL directorate is the lead on the program. Such a variable may capture organizational/managerial/technological differences. For this dataset, the TD variable is either RH or RQ. (This limitation is due to the sample size of the other TDs as previously discussed.) The performance type represents the partnership method between AFRL and the contractor. This variable consists of Research, Development, Test & Evaluation (RDT&E) and Small Business Innovative Research (SBIR) relationships. This type of variable may capture differences due to the size, skills, or knowledge of the company types (e.g., small versus large companies). TRL data for the S&T programs are used in seven different categorical variables. TRL Increase indicates if the TRL increases at any point during the program's lifecycle. Last Known TRL ≥ 6 denotes the last reported TRL of the program, while Final TRL ≥ 6 only analyzes programs that have a Final Research Summary Report. The decision to categorize based on TRL level 6 is due to the role this TRL level fulfills in the defense acquisition process. Specifically, a TRL of 6 is equivalent to demonstration in a relevant environment which is needed for a program to enter Milestone B (DoD, 2011). Four variables were created by grouping TRLs based on the maturity of the technology and the product's requirements, as determined in the literature (GAO, 1999). See Figure 2.

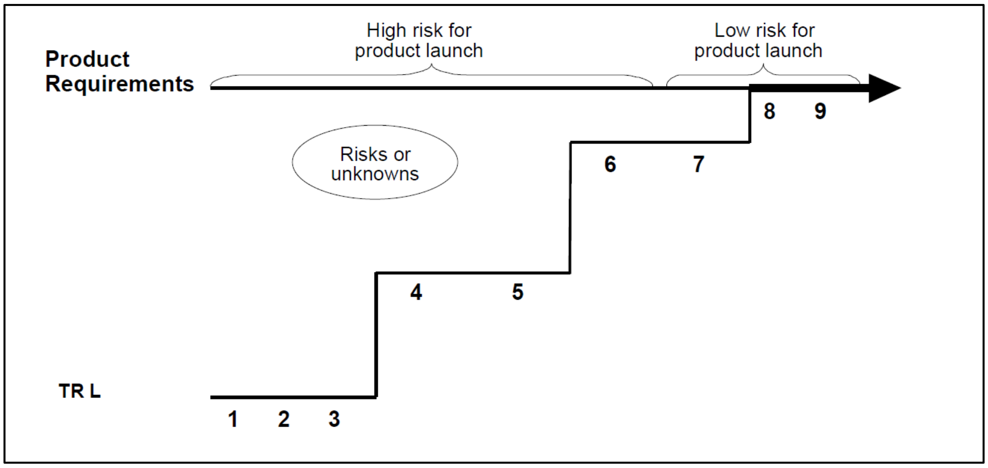

Figure 2. Using TRLs to Match Technology With Requirements
(GAO, 1999)

Additional variables of interest created from the Research Summary Report contract information include schedule growth, cost growth, and contract value. These attributes are commonly studied for acquisition programs at all phases of their lifecycles.

The variables for cost growth, schedule growth, and contract value have been converted from continuous variables to categorical variables in order to be included in the contingency table analysis. Binary (or dummy) variables with methodical break points were created in order to test the relationships at different locations. These breakpoints were derived either from the literature review or from descriptive statistics of the variable itself in the dataset with its mean and/or median. For example, the mean cost growth of the dataset was 68%, which led to the creation of a dummy variable (Cost Growth > 68%), separating programs that are above and below the sample mean. Likewise, Bolten et al. (2008) distinguished mean and median percentages of total DoD and Air Force acquisition program development cost percentages. These thresholds from Bolten et al. (2008) are also examined. A summary of the break points can be seen in Table 4.

Table 4. Break Point Summary

Category	Break Point	Reason	Source
Schedule Growth	0%	Any growth	Dataset
	33%	Median	Dataset
	63%	Mean	Dataset
Cost Growth	0%	Any growth	Dataset
	33.7%	DoD Development – Median	Bolten et al. (2008)
	44.1%	Air Force Development - Median	Bolten et al. (2008)
	56.5%	DoD Development – Mean	Bolten et al. (2008)
	60.5%	Air Force Development - Mean	Bolten et al. (2008)
	68%	Mean	Dataset
Contract Value	$1M	Median	Dataset
	$3M	Mean	Dataset

Results and Discussion

The contingency table results are organized into four sections: technical directorate, performance type, TRL, and growth relationships. Using the chi-square distribution as the test statistic, relationships are identified when Pearson's chi-squared test is significant at a p-value of less than 0.10. For highly significant results (p-value < 0.01), the odds ratio and its associated confidence interval are analyzed. It is important to note the possibility of spurious relationships. Spurious relationships occur when the two variables are associated but not causally related, possibly due to an unknown mediating variable. With the sheer number of 2 × 2 tables generated in this analysis, spurious relationships are possible. Therefore, only highly statistically significant results (p-value < 0.01) will be studied in detail (i.e., full contingency table shown), while the other significant variables are observed solely as potential findings.

Technical Directorate

The TD categorical variable denotes which AFRL directorate is the lead on the respective program: either RH (Airman Systems) or RQ (Aerospace Systems). Analyzing the TD variable resulted in 19 contingency tables to be tested for significance. Two variables were significant at an alpha of 0.10, and two were significant at an alpha of 0.05. The full set of test results is provided in Table 5.

Table 5. Contingency Table Results for Technical Directorate

Variable	TD
Performance Type	
TRL Increase	**
Last Known TRL ≥ 6	
Final TRL ≥ 6	
TRL 1-3	
TRL 4-5	
TRL 6-7	
TRL 8-9	
Schedule Growth > 0%	
Schedule Growth > 33% (Median)	**
Schedule Growth > 63% (Mean)	*
Contract Value > $1.0M (Median)	
Contract Value > $3.0M (Mean)	
Cost Growth > 0%	*
Cost Growth > 33.7% (DoD Dev - Median)	
Cost Growth > 44.1% (AF Dev - Median)	
Cost Growth > 56.5% (DoD Dev - Mean)	
Cost Growth > 60.5% (AF Dev - Mean)	
Cost Growth > 68% (Mean)	
Total Significant Contingency Tables:	**4**

Table Legend:
* p-value < 0.10
** p-value < 0.05
*** p-value < 0.01

TRL Increase is the only TRL variable with a statistically significant relationship to TD. This test suggests that it is more probable to have a program's TRL increase with RQ (Aerospace Systems) programs than with RH (Airman/Human Systems) programs. The RQ programs are comprised primarily of engine and propulsion (hardware) system technologies. The ability to transition RQ through TRL levels may be due to the relationship of hardware versus software (human systems interactions). It is likely easier to make advancements in hardware technologies as the testing, failures, and efficiencies may be more conclusive.

Similarly, the contingency table results suggest that RQ programs are more probable to have cost growth as well as schedule growth that is greater than 33% (the dataset's median) and 63% (the dataset's mean). This could be related to the maturing technology (increasing the TRL) of RQ programs. If the technology is maturing, a program office may be more likely to increase funding and schedule to keep the maturation on track. If the technologies do not mature, it could be that the agile nature of S&T programs allows for early decisions to cancel programs. In summary, the TD results suggest that RQ programs are more likely to technologically mature, have cost growth, and have schedule growth (greater than the dataset mean and median) when compared to RH programs.

Performance Type

The performance type variable represents the partnership method between AFRL and the contractor: either Research, Development, Test & Evaluation (RDT&E) or Small Business Innovative Research (SBIR) relationships. This variable formed 19 contingency tables to be tested for significance. One variable was significant at an alpha of 0.10, two variables were significant at an alpha of 0.05, and two variables were significant at an alpha of 0.01. The full set of test results is provided in Table 6.

Table 6. Contingency Table Results for Performance Type

Variable	Performance Type
TD	
TRL Increase	
Last Known TRL ≥ 6	**
Final TRL ≥ 6	**
TRL 1-3	
TRL 4-5	
TRL 6-7	
TRL 8-9	
Schedule Growth > 0%	*
Schedule Growth > 33% (Median)	
Schedule Growth > 63% (Mean)	
Contract Value > $1.0M (Median)	***
Contract Value > $3.0M (Mean)	***
Cost Growth > 0%	
Cost Growth > 33.7% (DoD Dev - Median)	
Cost Growth > 44.1% (AF Dev - Median)	
Cost Growth > 56.5% (DoD Dev - Mean)	
Cost Growth > 60.5% (AF Dev - Mean)	
Cost Growth > 68% (Mean)	
Total Significant Contingency Tables:	**5**

Table Legend:
* p-value < 0.10
** p-value < 0.05
*** p-value < 0.01

Table 6 test results suggest that an S&T program with an RDT&E performance type is more likely to have or end with a TRL of at least 6 than an SBIR type program is. SBIR programs are developed by small domestic businesses, which potentially provides an agile way to stimulate high-tech innovation. But RDT&E programs are dominated by the larger, more experienced defense contractors. These results suggest that the larger defense contractors may obtain contracts with more mature technologies due to their capacity and ability to develop these technologies when compared to SBIR businesses.

Furthermore, as a potential indication of RDT&E and SBIR working different kinds of programs from the start, one can observe that it is more probable to have contract values greater than $1 million (the dataset's median) with RDT&E performance types, as seen in Figure 3. Testing significance when the contract value is greater than $3 million (the dataset's mean) produces similar results to Figure 3, with an even smaller p-value. Again, this could be due to the differences in the types of contractors involved in RDT&E and SBIR programs. Larger defense contractors possibly obtain larger programs because they have more breadth of experience or capacity, while the small businesses obtain smaller contracts with a more constrained objective; the acquisition community often sees a similar relationship when the large defense contractors are prime on a large system and smaller vendors are subcontractors for a particular subsystem. Additionally, SBIR programs may target uncertain and risky technologies that small businesses research so that AFRL can evaluate which programs have the potential to develop into mature technologies. The scale of these uncertain programs may contribute to lower contract values. In fact, the odds ratio indicates that given the program has an SBIR performance type, the odds of the contract value being less than $1 million is 9.7 times higher than when the program has an RDT&E performance type.

Contingency Table

Last Contract Value>$1M (Med)

Performance Type: Count / Expected	0	1	Total
RDTE	3	12	15
	7.69231	7.30769	
SBIR	17	7	24
	12.3077	11.6923	
Total	20	19	39

Tests

N	DF	-LogLike	RSquare (U)
39	1	5.0265877	0.1860

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	10.053	0.0015*
Pearson	9.547	0.0020*

Odds Ratio

Odds Ratio	Lower 95%	Upper 95%
0.102941	0.022042	0.480764

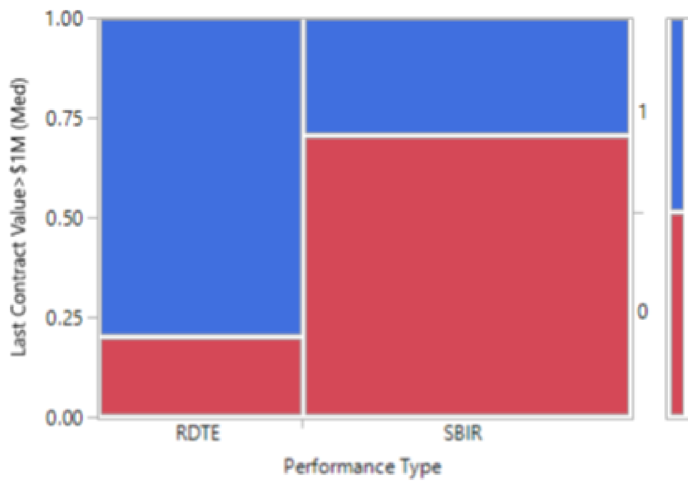

Figure 3. Contingency Table of Performance Type by Contract Value > $1 Million

The Table 6 contingency test results also suggest that a program with an SBIR performance type is more likely to have schedule growth. While test results indicate that RDT&E programs are more likely to have higher TRL levels, the opposite could be said in that SBIR programs are more likely to have lower TRL levels. Less is known about these immature technologies, which could lead these small businesses to spend more time developing them, leading to schedule slippage. This result is consistent with the literature findings of Dubos et al. (2008).

In summary, the results suggest that a program that has a performance type of RDT&E is more likely to have a TRL of 6. Furthermore, highly significant results point to evidence that a program that has a performance type of RDT&E is more likely to have a contract value greater than $1 million. Lastly, the results suggest that SBIR programs are more likely to experience schedule growth.

Technology Readiness Level

TRL data was used to create seven different binary variables, as previously discussed. These seven TRL variables were tested for significance against the 11 performance variables to produce 77 contingency tables. Seven variables were significant at an alpha of 0.10, four variables were significant at an alpha of 0.05, and one variable was significant at an alpha of 0.01. Despite registering significant Pearson p-values, the contingency table results for the seven significant variables at an alpha of 0.10 were found to be invalid. For all seven tests, the expected counts of two of the four cells were less than five. This violates an assumption for a valid chi-squared contingency table test, which states the sample size should be large enough so that the estimated expected count will be equal to five or more. As a further check, Fisher's Exact Test—which is a non-parametric test for small samples—found all seven tests to be non-significant. This result was largely due to the small number of programs with a TRL of 6–7 (5) and a Final TRL of ≥ 6 (4). The full set of test results is provided in Table 7 with special subscript designators on those test results deemed invalid.

Table 7. Significant Contingency Tables for Technology Readiness Level

Variable	TRL Increase	Last Known TRL ≥ 6	Final TRL ≥ 6	TRL 1-3	TRL 4-5	TRL 6-7	TRL 8-9
Schedule Growth > 0%		**	$*_1$			$*_1$	
Schedule Growth > 33% (Median)							
Schedule Growth > 63% (Mean)							
Contract Value > $1.0M (Median)				**			
Contract Value > $3.0M (Mean)		**				***	
Cost Growth > 0%						$*_1$	
Cost Growth > 33.7% (DoD Dev - Median)						$*_1$	
Cost Growth > 44.1% (AF Dev - Median)						$*_1$	
Cost Growth > 56.5% (DoD Dev - Mean)						$*_1$	
Cost Growth > 60.5% (AF Dev - Mean)						$*_1$	
Cost Growth > 68% (Mean)						**	
Total Significant Contingency Tables:	**0**	**2**	**1**	**1**	**0**	**8**	**0**

Table Legend:
- $*_1$ p-value < 0.10, 50% of Expected Counts < 5, Non-significant Fisher's Exact Test
- * p-value < 0.10
- ** p-value < 0.05
- *** p-value < 0.01

The contingency table results suggest that an S&T program is *more* likely to have cost growth greater than 68% (the dataset's mean) with a TRL of 6 or 7 but *less* likely to have schedule growth with a TRL ≥ 6. Such a finding, perhaps unusual for a development program, is both intuitive and precedent in an S&T context. With an early TRL (1–5), there is little knowledge of how the technology will mature. This poses a problem to program managers and cost estimators. As technologies mature, investments are made, which allow costs to grow over their initial estimates. As the technology integrates into a demonstration effort (TRL 6–8), the program is often met with new and unexpected challenges, which tends to increase costs.

These results support previous literature conducted on Air Force programs which concluded that estimated costs vary exponentially across time with the progression through the various TRLs (Smoker & Smith, 2007). However, for more mature technologies, there is a broader knowledge base available for the technology's development due to more completed research. With a higher TRL, and thus more knowledge of the technology available, the better the chance of meeting schedule requirements (Dubos et al., 2008). This literature finding is also consistent with the results found here.

Table 7 results also suggest that an S&T program is *more* likely to have contract values greater than $3 million (the dataset's mean) with a TRL of 6 or greater and *less* likely to have contract values greater than $1 million (the dataset's median) with a TRL of 1–3. The explanation is consistent with the aforementioned cost growth finding. As the program's technology matures, additional investments are made, as shown in the contingency analysis results in Figure 4. In fact, the odds ratio indicates that given the program has a TRL of 6 or 7, the odds of the contract value being greater than $3 million is 14.5 times higher than a program with a TRL other than 6 or 7.

Contingency Table

Last Contract Value>$3M (Mean)

TRL 6-7	Count / Expected	0	1	Total
0	Count	29	8	37
	Expected	26.4286	10.5714	
1	Count	1	4	5
	Expected	3.57143	1.42857	
Total		30	12	42

Tests

N	DF	-LogLike	RSquare (U)
42	1	3.3084592	0.1317

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	6.617	0.0101*
Pearson	7.356	0.0067*

Odds Ratio

Odds Ratio	Lower 95%	Upper 95%
14.5	1.415167	148.5691

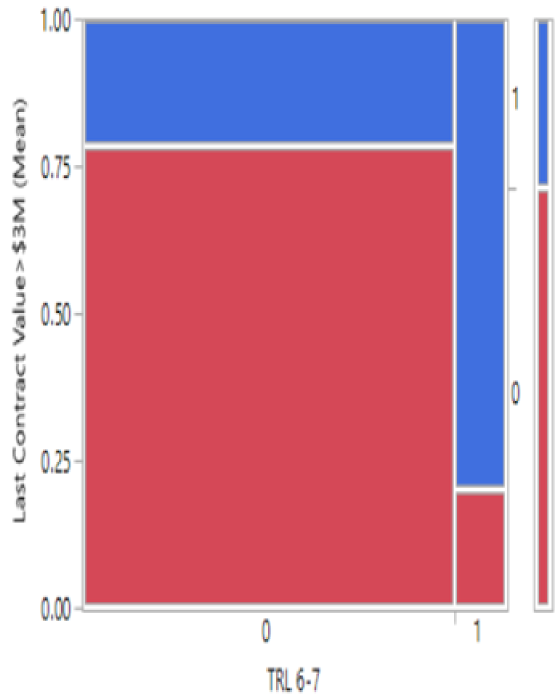

Figure 4. Contingency Table of TRL 6–7 by Contract Value > $3 Million

In summary, the TRL results suggest that programs with mature technologies are more likely to experience larger than average cost growth and larger contract values. Additionally, these programs are less likely to experience schedule growth. Furthermore, the results suggest that programs with immature technologies are less likely to have larger contract values.

Growth Relationships

As previously shown, variables for TD, performance type, and TRL were tested for their relationships with cost growth, schedule growth, and contract value. An analysis was also conducted among these latter variables to analyze their relationships to each other; a total of 63 relationships were tested for significance. Eight tests were significant at an alpha of 0.10, 11 tests were significant at an alpha of 0.05, and 22 tests were significant at an alpha of 0.01. The full set of test results is provided in Table 8.

Table 8. Significant Contingency Tables for Growth Relationships

Variable	Schedule Growth > 0%	Schedule Growth > 33% (Med)	Schedule Growth > 63% (Mean)	Contract Value > $0.9M	Contract Value > $1.0M (Med)	Contract Value > $3.0M (Mean)	Contract Value > $4.0M	Contract Value > $5.0M	Total Significant Cont. Tables
Contract Value > $0.9M		**	**						2
Contract Value > $1.0M (Median)									0
Contract Value > $3.0M (Mean)									0
Contract Value > $4.0M									0
Contract Value > $5.0M									0
Cost Growth > 0%	**	***	***	***	***	**	**	*	8
Cost Growth > 33.7% (DoD Dev - Median)	*	*	***		***	***	***	**	7
Cost Growth > 44.1% (AF Dev - Median)	*	*	***		***	***	***	**	7
Cost Growth > 56.5% (DoD Dev - Mean)	*	**	***			***	***	**	6
Cost Growth > 60.5% (AF Dev - Mean)	*	**	***			***	***	**	6
Cost Growth > 68% (Mean)		*	***			***	***	***	5
Total Significant Contingency Tables:	**5**	**7**	**7**	**1**	**3**	**6**	**6**	**6**	**41**

Table Legend:
- * p-value < 0.10
- ** p-value < 0.05
- *** p-value < 0.01

The contingency table results suggest that it is more probable for S&T programs with larger contract values to experience cost growth. Observing cost growth relationships with the original two contract value variables (using the mean and median of the dataset) provided highly significant results. To explore the sensitivity of these relationships relative to the threshold used to define the binary variables, additional contract value variables were created with lower and higher breakpoints. This additional analysis found contract values greater than $0.9 million to be the lowest threshold for which a statistically significant relationship could be found with amount of cost growth (i.e., cost growth > 0%). As the contract value threshold increased, additional cost growth variables displayed statistical significance until all were significant at a contract value of $3 million. This suggests that cost growth and contract value have a positive correlation with each other.

Table 8 results also suggest that it is more probable for S&T programs with contract values greater than $0.9 million to experience schedule growth above the median and mean (i.e., greater than 33% and 63%, respectively). This was the only contract value variable to result in significant p-values when tested with schedule growth variables. These results imply that programs with contract values less than $0.9 million are less likely to experience schedule growth.

Finally, the results suggest that if S&T programs are experiencing schedule growth, then it is more likely that they're also experiencing cost growth. This seems to contradict the findings that programs with mature technologies are more likely to experience cost growth while being less likely to experience schedule growth. But further analysis of these results suggests that programs with large schedule growth percentages are even more likely to experience cost growth at all amounts. This is because it is the immature technology programs that are experiencing both the schedule and cost growth.

In summary, the results suggest that S&T programs with larger contract values experience cost growth, while programs with smaller contract values are less likely to experience schedule growth. Finally, analyzing the relationship between cost and schedule growth suggests that programs with schedule growth are more likely to have cost growth as well. Deeper analysis revealed that this schedule growth/cost growth relationship is found in those programs with immature technologies.

Conclusion

S&T programs serve an important role in the defense acquisition process. They constitute the initial phase of the acquisition process through discovery and development of warfighting technology. The results of these programs impact the larger MDAPs that integrate the technologies in subsequent phases of the lifecycle. Despite this important role, little prior research has examined the performance of S&T programs. Thus, the overarching goal of this paper was to discern new insights from an analysis of S&T program characteristics in relation to their program's performance.

The literature review identified technological maturity as a critical success factor in product development (GAO, 1999). One measure defense programs use for technological maturity is TRL levels. TRLs, therefore, were an integral component under investigation in this analysis. The objective was to understand how TRLs affect S&T program performance. There are several key findings.

First, the results suggest that aerospace programs are more likely to technologically mature when compared to human system programs. In other words, the AFRL aerospace programs are more likely to increase the TRLs in their programs. To the extent that technological maturity is a measure of success, the aerospace programs outperform. However, this technical performance comes at a cost, as the aerospace programs were also more likely to experience cost and schedule growth. Intuitively, these results are compatible; with proven success in technology maturation, increases in funding and schedule are likely to keep the maturation on track.

Second, the partnership method between the government and contractor matters. The partnerships for S&T programs consist of SBIR and RDT&E relationships. The RDT&E programs are more likely to have and end with a TRL of 6 or more in comparison to SBIR programs. The result is not entirely surprising because, by definition, the larger defense companies comprise the RDT&E category. These larger companies have the capacity and resources to mature technology that the smaller SBIR companies may not possess.

Third, TRLs and program performance are linked. The relationships with TRLs suggest that programs with mature technologies are more likely to experience above-average cost growth and larger contract values while less likely to experience schedule growth. Additionally, the results suggest that programs with immature technologies are less likely to have larger contract values. As technologies mature, additional funds for investments are made, which increases costs over their initial contract values. This is likely to happen when the program is met with new and unexpected challenges as the technology integrates into a demonstration effort (TRL 6–8). Linick (2017) found that as the TRL increased throughout the development phase, the percentage of the development cost increased at an increasing rate. This literature finding is in agreement with these results. Conversely, as these technologies mature, there is a broader knowledge base for their development, which increases the chance of meeting schedule requirements.

Lastly, the analysis of “growth” variables (cost growth, schedule growth, and contract value) provides additional insights on S&T programs. Specifically, the analysis suggests that S&T programs with larger contract values experience larger cost growth at the same time programs with smaller contract values are less likely to experience schedule growth. Further analyzing the relationship between cost and schedule growth, the results suggest that if programs have larger schedule growth, then they are more likely to have larger cost growth as well. Deeper analysis revealed that this schedule growth/cost growth relationship is found in those programs with immature technologies.

Prior examinations of S&T programs are scarce. Thus, the possibilities for future research are vast. The exploratory analysis conducted here focused solely on AFRL programs. S&T programs in the other military services warrant examination. Additionally, one of the more surprising aspects of the data obtained from S&T programs was the reported TRL at various stages of the program’s lifecycle. In order for a program to advance past Milestone B into the EMD phase, a program must have a TRL of 6 or greater. Further research into those S&T programs whose technology matured (TRL increased) could identify common characteristics which indicate a higher probability of technological maturation. The exploratory analysis provided here was just the first step of the journey. Through future research and discoveries, the knowledge needed to increase the odds for successful S&T programs is possible.

Disclaimer: The views expressed in this article are those of the authors and do not reflect the official policy or position of the U.S. Air Force, Department of Defense, or the U.S. government.

Appendix A. TRL Definitions, Descriptions, and Supporting Information

TRL	Definition	Description	Supporting Information
1	Basic principles observed and reported.	Lowest level of technology readiness. Scientific research begins to be translated into applied research and development (R&D). Examples might include paper studies of a technology’s basic properties.	Published research that identifies the principles that underlie this technology. References to who, where, when.
2	Technology concept and/or	Invention begins. Once basic principles are observed, practical	Publications or other references that outline the application

TRL	Definition	Description	Supporting Information
	application formulated.	applications can be invented. Applications are speculative, and there may be no proof or detailed analysis to support the assumptions. Examples are limited to analytic studies.	being considered and that provide analysis to support the concept.
3	Analytical and experimental critical function and/or characteristic proof of concept.	Active R&D is initiated. This includes analytical studies and laboratory studies to physically validate the analytical predictions of separate elements of the technology. Examples include components that are not yet integrated or representative.	Results of laboratory tests performed to measure parameters of interest and comparison to analytical predictions for critical subsystems. References to who, where, and when these tests and comparisons were performed.
4	Component and/or breadboard validation in a laboratory environment.	Basic technological components are integrated to establish that they will work together. This is relatively "low fidelity" compared with the eventual system. Examples include integration of "ad hoc" hardware in the laboratory.	System concepts that have been considered and results from testing laboratory scale breadboard(s). References to who did this work and when. Provide an estimate of how breadboard hardware and test results differ from the expected system goals.
5	Component and/or breadboard validation in a relevant environment.	Fidelity of breadboard technology increases significantly. The basic technological components are integrated with reasonably realistic supporting elements so they can be tested in a simulated environment. Examples include "high-fidelity" laboratory integration of components.	Results from testing laboratory breadboard system are integrated with other supporting elements in a simulated operational environment. How does the "relevant environment" differ from the expected operational environment? How do the test results compare with expectations? What problems, if any, were encountered? Was the breadboard system refined to more nearly match the expected system goals?
6	System/subsystem model or prototype demonstration in a relevant environment.	Representative model or prototype system, which is well beyond that of TRL 5, is tested in a relevant environment. Represents a major step up in a technology's demonstrated readiness. Examples include testing a prototype in a high-fidelity laboratory environment or in a simulated operational environment.	Results from laboratory testing of a prototype system that is near the desired configuration in terms of performance, weight, and volume. How did the test environment differ from the operational environment? Who performed the tests? How did the test compare with

TRL	Definition	Description	Supporting Information
			expectations? What problems, if any, were encountered? What are/were the plans, options, or actions to resolve problems before moving to the next level?
7	System prototype demonstration in an operational environment.	Prototype near or at planned operational system. Represents a major step up from TRL 6 by requiring demonstration of an actual system prototype in an operational environment (e.g., in an aircraft, in a vehicle, or in space).	Results from testing a prototype system in an operational environment. Who performed the tests? How did the test compare with expectations? What problems, if any, were encountered? What are/were the plans, options, or actions to resolve problems before moving to the next level?
8	Actual system completed and qualified through test and demonstration.	Technology has been proven to work in its final form and under expected conditions. In almost all cases, this TRL represents the end of true system development. Examples include developmental test and evaluation (DT&E) of the system in its intended weapon system to determine if it meets design specifications.	Results of testing the system in its final configuration under the expected range of environmental conditions in which it will be expected to operate. Assessment of whether it will meet its operational requirements. What problems, if any, were encountered? What are/were the plans, options, or actions to resolve problems before finalizing the design?
9	Actual system proven through successful mission operations.	Actual application of the technology in its final form and under mission conditions, such as those encountered in operational test and evaluation (OT&E). Examples include using the system under operational mission conditions.	OT&E reports.

Appendix B. Sample Research Summary Report

Research Summary Report

General Information

Work Unit Title:	Program Title		
Work Unit (WU) #:		Accession #:	XXXXX
Start Date:	MM/DD/YYYY	End Date:	MM/DD/YYYY
Lead TD:	RQ	Effort Security:	Unclassified
Literature Search Order #:	XXXXX	Lit Search Date:	MM/DD/YYYY
Work Phase Code:		WU Category:	
Responsible Organization:	Aerospace Systems Directorate (AFRL/RQ) WRIGHT-PATTERSON AFB, OH		
Parent Program:	Parent Program Title		

Work Unit Manager

Name:	XXXX XXXX	Phone:	(XXX) XXX-XXXX
Office Symbol:	RQTE	Email:	xxxx.xxxx@us.af.mil

DoD Required Information

Performance Method:	PROCUREMENT/ACQUISITION AWARD - Contract,	Lab Core Technical Competency:	Power, Propulsion, Energy & Alternative Fuels
Performance Type:	RDTE - Research, Development, Test, & Evaluation Work Unit	Technology Readiness Level:	2 Technology Concept
Fields of Science & Engineering:		Data Management Plan Exist:	No
Joint Capability Area:	Force Application		
Communities of Interest:	1.1 Aircraft Propulsion, Power and Thermal		
Technology Transition Opportunities:	None at this time. Follow on contract will progress TRL of key components to enable future transition.		

Principal Investigator

Name:	XXXX XXXXXX	Phone:	XXX-XXX-XXXX
Office Symbol:	XX	Email:	xxxx.xxxxx@xx.com

Performing Mechanism

Contract #:	FAXXXX-XX-X-XXXX-XXXX	Contract Status:	Complete
Contract Face Value	$XX,XXX,XXX	TR Due DTIC Date:	MM/DD/YYYY
Award Date:	MM/DD/YYYY	TR Draft Due Date:	MM/DD/YYYY
Tech End Date:	MM/DD/YYYY	Terminate Date:	
End Date:	MM/DD/YYYY	WAWF /DD250Z Date:	MM/DD/YYYY

Performing Organization

Performer:	NAME OF COMPANY	Address:
Division:	XX	
CAGE Code:	XXXXXX	
Performing Org Type:	Industrial Firm	

WU Funding ($K)

PE / BPAC	Source	Funding	Prior	FY18	FY19	FY20	FY21	FY22	FY23	To Complete	Total
		Appropriated									
		Obligations									
		Appropriated									
		Obligations									

Descriptive and Progress/Status

Objective/Restricted

MMM DD, YYYY
12:00:00 AM

Distribution C US Government Agencies and their Contractors only - C-CL Critical Technology (Export Controlled)
Page 1

References

Bolten, J. G., Leonard, R. S., Arena, M. V., Younossi, O., & Sollinger, J. M. (2008). *Sources of weapon system cost growth* (Report No. MG-670). RAND.

Cancian, M. F. (2010). Cost growth: Perception and reality. *Defense Acquisition Research Journal*, *17*(3), 389–404.

DoD. (2009, July). *Technology readiness assessment (TRA) deskbook*. Office of the Director, Defense Research and Engineering.

DoD. (2011, May 13). *Technology readiness assessment (TRA) guidance.* Assistant Secretary of Defense for Research and Engineering.

Duffner, R. (2000). *Science and technology: The making of the Air Force research laboratory*. Air University Press.

Dubos, G., Saleh, J., & Braun, R. (2008). Technology readiness level, schedule risk and slippage in spacecraft design: Data analysis and modeling. *Journal of Spacecraft and Rockets*, *45*(4), 836–842. https://doi.org/10.2514/6.2007-6020

Gorn, M. H. (1988). *Harnessing the genie: Science and technological forecasting and the Air Force 1944–1986*. Office of Air Force History.

GAO. (1999). *Best practices: Better management of technology development can improve weapon systems outcomes* (GAO/NSIAD-99-162).

Katz, D. R., Sarkani, S., Mazzuchi, T., & Conrow, E. H. (2015). The relationship of technology and design maturity to DoD weapon system cost change and schedule change during engineering and manufacturing development. *Systems Engineering*, *18*(1), 1–15. https://doi.org/10.1111/sys.21281

Linick, J. (2017, June 6–9). *Technology readiness level (TRL) vs. percent development cost* [Presentation]. ICEAA Professional Development & Training Workshop, Portland, OR, United States.

Meredith, J., & Mantel, S. (2003). *Project management: A managerial approach*. John Wiley & Sons.

Nasir, M. H. N., & Sahibuddin, S. (2011). Critical success factors for software projects: A comparative study. *Scientific Research and Essays*, *6*(10), 2174–2186. https://doi.org/10.5897/SRE10.1171

Office of the Chief Scientist of the U.S. Air Force. (2010). *Technology horizons: A vision for Air Force science & technology During 2010–2030*. Department of the Air Force.

Pinto, J. K., & Slevin, D. P. (1987). Critical factors in successful project implementation. *IEEE Transactions on Engineering Management*, *34*(1), 22–27.

Pinto, J. K., & Slevin, D. P. (1998, September). Project success: Definitions and measurement techniques. *Project Management Journal*, *19*(3), 67–73.

Rendon, R. (2012). Defense procurement: An empirical analysis of critical success factors. In *Charting a course in public procurement innovation and knowledge sharing* (pp. 174–208).

Rodriguez-Segura, E., Ortiz-Marcos, I., Romero, J. J., & Tafur-Segura, J. (2016). Critical success factors in large projects in the aerospace and defense sectors. *Journal of Business Research*, *69*(11), 5419–5425.

Sadin, S. R., Povinelli, F. P., & Rosen, R. (1989). The NASA technology push towards future space mission systems. *Acta Astronautica*, *20*, 73–79. https://doi.org/10.1016/0094-5765(89)90054-4

Smirnoff, J. P., & Hicks, M. J. (2008). The impact of economic factors and acquisition reforms on the cost of defense weapon systems. *Review of Financial Economics*, *17*(1), 3–13.

Smoker, R. E., & Smith, S. (2007). System cost growth associated with technology-readiness level. *Journal of Parametrics*, *26*(1), 8–38. https://doi.org/10.1080/10157891.2007.10462276

Tishler, A., Dvir, D., Shenhar, A., & Lipovetsky, S. (1996). Identifying critical success factors in defense development projects: A multivariate analysis. *Technological Forecasting and Social Change*, *51*(2), 151–171. https://doi.org/10.1016/0040-1625(95)00197-2

Tracy, S. P., & White, E. D. (2011). Estimating the final cost of a DoD acquisition contract. *Journal of Public Procurement*, *11*(2), 190–205. https://doi.org/10.1108/jopp-11-02-2011-b002

U.S. Air Force. (2019). *Science and technology strategy: Strengthening USAF science and technology for 2030 and beyond*. U.S. Government Printing Office.

Zwikael, O., & Globerson, S. (2006). From critical success factors to critical success processes. *International Journal of Production Research*, *44*(17), 3433–3449.