NPS-TE-23-203

ACQUISITION RESEARCH PROGRAM
SPONSORED REPORT SERIES

Operational Test in Agile Development: Vetted Capability at the Speed of Relevance

June 2023

LT Brian D. Morrill, USN
LCDR Joseph D. Polnaszek, USN

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

Department of Defense Management

Naval Postgraduate School

Prepared for the Naval Postgraduate School, Monterey, CA 93943

ACQUISITION RESEARCH PROGRAM
DEPARTMENT OF DEFENSE MANAGEMENT
NAVAL POSTGRADUATE SCHOOL

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

To request defense acquisition research, to become a research sponsor, or to print additional copies of reports, please contact the Acquisition Research Program (ARP) via email, arp@nps.edu or at 831-656-3793.

Acquisition Research Program
Department of defense management
Naval Postgraduate School

ABSTRACT

The traditional Operational Test and Evaluation (OT&E) process and timeline are no longer sufficient to keep pace within an agile software development environment. Agile development requires OT&E to be just as agile while still ensuring the adequacy and sufficiency required to keep the Fleet's trust in new capabilities.

In order to accomplish this, we recommend five initiatives that will help the current OT&E process. Those initiatives are: The "Pulverizer," an agile Test and Evaluation Master Plan (TEMP), universal test plan working group, an integrated operational test assessment, and a dedicated annual OT&E period. These initiatives make the bureaucratic requirements for OT&E much more streamlined and agile. They ensure an annual schedule all stakeholders can work towards and drive down many of the important OT&E decisions to the test squadron level.

These recommendations have already been shown to work with positive feedback from stakeholders for the F/A-18 program. If these changes are implemented across all of naval aviation, OT&E can become as agile as the software development.

THIS PAGE INTENTIONALLY LEFT BLANK

ABOUT THE AUTHORS

Lieutenant Brian "Bobby Hill" Morrill is a native of Salt Lake City, Utah and graduated from Utah Valley University with a Bachelor of Science in Emergency Services. After attending Officer Candidate School in Newport, Rhode Island he was commissioned an Ensign in October 2014. He has a Master's of Science degree in Program Management from the Naval Postgraduate School.

LT Morrill's tours were with Carrier Airborne Early Warning Squadron ONE TWO SIX (VAW-126), where he flew combat mission in support of OPERATION INHERENT RESOLVE. From June 2020 to May of 2023 he was attached to AIRTEVRON ONE (VX-1) where he served as Assistant Operational Test Director in support of testing of new hardware and software for the E-2D. LT Morrill is currently attached to US Naval Forces Central Command where he is flag aide for the Deputy Commander.

LT Morrill has been constantly supported by his beautiful wife Shanna, and 3 children Isabelle (6), Alexis (5), and Jack (3).

LCDR Joseph "Tommy Boy" Polnaszek is a native of Middletown, Rhode Island. He is a 2007 graduate of Boston University where he earned a Bachelor of Science in Electrical Engineering. He was designated a Naval Flight Officer in 2010. He is currently studying for a Master's of Science in Program Management from the Naval Postgraduate School in Monterey, California.

LCDR Polnaszek's operational tours at Naval Air Station (NAS) Lemoore, California include Strike Fighter Squadron 154 (VFA-154) and Carrier Air Wing Nine (CVW-9) where he served as the Strike Operations Officer. He flew in Operations ENDURING FREEDOM, FREEDOM'S SENTINEL, and INHERENT RESOLVE, while deployed on USS Nimitz (CVN 68) and USS John C. Stennis (CVN 74).

LCDR Polnaszek, also completed two shore tours at Air Test and Evaluation Squadron Nine (VX-9) as the Operational Test Director of System Configuration Set

H12 and Operations Officer.

LCDR Polnaszek has logged over 1,700 flight hours in the F/A-18F, EA-18G, T-45, T-39, T-1A, and T-6 and has accumulated over 300 carrier arrested landings.

His personal decorations include the Strike/Flight Air Medal, Navy Commendation Medal (2 awards), And Navy and Marine Corps Achievement Medal, and various other personal and unit awards.

LCDR Polnaszek's service is made possible by his beautiful wife Caitlin

NPS-TE-23-203

ACQUISITION RESEARCH PROGRAM
SPONSORED REPORT SERIES

Operational Test in Agile Development: Vetted Capability at the Speed ff Relevance

June 2023

LT Brian D. Morrill, USN
LCDR Joseph D. Polnaszek, USN

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

Department of Defense Management

Naval Postgraduate School

Prepared for the Naval Postgraduate School, Monterey, CA 93943

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AAF	Adaptive Acquisition Framework
ACAT	Acquisition Category
AI	Artificial Intelligence
AIRTEVRON	Air Test and Evaluation Squadron
AWL	Advanced Weapons Lab
CNAL	Commander, Naval Air Forces
COI	Critical Operational Issue
COP	Common Operational Picture
DA	Decision Authorities
DAOT&E	Dedicated Annual Operational Test and Evaluation
DAS	Defense Acquisition System
DoD	Department of Defense
DoDI	Department of Defense Instruction
DOT&E	Director, Operational Test and Evaluation
DSSC	Delta System Software Configuration
DT	Developmental Test
EA	Electronic Attack
FTE	Flight Test Engineers
GAO	Government Accountability Office
IEF	Integration Evaluation Framework
IOTA	Initial Operational Test Assessment
ISIC	Immediate Superior in Chain of Command
IT	Integrated Test
JCIDS	Joint Capabilities Integration and Development Systems
LFE	Large Force Exercise
LTE	Lead Test Engineer
MBTD	Mission-Based Test Design

MDA	Milestone Decision Authority
MDPA	Major Defense Acquisition Program
MIDS	Multifunctional Information Display System
MVCR	Minimum Viable Capability Release
MVP	Minimum Viable Product
NAWCWD	Naval Air Warfighting Center-Weapons Division
OEM	Original Equipment Manufacturer
OPTEVFOR	Operational Test and Evaluation Force
OT	Operational Test
OTA	Operational Test Authority
OTD	Operational Test Director
OTG	Operational Tactics Guide
OT&E	Operational Test and Evaluation
OTRR	Operational Test Readiness Review
PEO	Program Executive Offices
PM	Program Manager
POMs	Program Objective Memorandums
RO	Requirements Officer
ROC/POE	Required Operational Capability/Projected Operational Environment
RCRM	Running Comment Resolution Matrix
SAFe	Scaled Agile Framework
SCS	Software Configuration Set
SECNAVINST	Secretary of the Navy Instruction
SME	Subject Matter Expert
SOR	Statement of Requirement
SUT	System Under Test
SWP	Software Acquisition Pathway
TACAN	Tactical Aid to Navigation
TEMP	Test and Evaluation Master Plan

TTP	Techniques, Tactics, and Procedures

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The United States acquisition process has undergone major changes in the last few years as their focus has shifted from non-state aggressor terrorist organization in the Middle East to near-peer threats in Asia with China and, to a lesser extent, in Europe with Russia. Both China and Russia have been fielding greater capabilities at alarming rates. In a Government Accountability Office (GAO) report from 1995 it was noted that "since 1989 the official Chinese defense budget increased annually by double digits" (GAO, 1995). However, the GAO did not seem too concerned about these increases noting that "to date few new weapon systems have been acquired and other improvements such as better training have only impacted a few units" (GAO, 1995).

Contrast this view with another GAO report from February of 2022 where they state that "China has turned what was once an obsolete military into one that can challenge the U.S. military across the spectrum of conventional and unconventional capabilities" (GAO, 2022). That is a huge change in less than 30 years. Today the U.S. military uses many of the same weapon systems that it used in 1995 including the Aegis Weapon System and the F18 Hornet. While these systems have been upgraded in this time, they are still leveraging primarily legacy technologies.

China on the other hand has new technologies that include artificial intelligence (AI), unmanned vehicles, and biotechnologies (Director of National Intelligence, 2021). To partially address this threat from an acquisition perspective, the Department of Defense, (DoD) has implemented the Adaptive Acquisition Framework (AAF) that allows Milestone Decision Authorities (MDA), other Decision Authorities (DA), and Program Managers (PM) to "develop acquisition strategies and employ acquisition processes that match the characteristics of the capability being acquired" (Under Secretary of Defense for Acquisition and Sustainment, 2020a).

One of the options available to the decision makers is the software acquisition pathway. The goal of this pathway is to provide a "rapid, iterative approach" to software development to "reduce costs, technological obsolescence, and acquisition risk" (Under Secretary of Defense for Acquisition and Sustainment, 2020a). The goal of this rapid and

iterative approach, also known as agile development, is meant to move software away from monolithic delivery of major weapon systems such as the AEGIS weapon system on U.S. Navy Destroyers and Cruisers or the F/A-18 Super Hornet Software Configuration Set (SCS) to smaller and more manageable software updates much like updates on a personal computer or application on a smart phone.

A. PROBLEM

While the U.S. Navy's Operational Test and Evaluation Force (OPTEVFOR) has acknowledged the need to implement agile software development, they do not have any polices in place for Operational Test and Evaluation (OT&E) squadrons to follow. The traditional OT&E processes and timelines are no longer sufficient to keep pace within an agile software development environment. Agile development requires OT&E to be just as agile while still ensuring the adequacy and sufficiency required to keep the Fleet's trust in new capabilities.

B. RESEARCH QUESTIONS

How can OT&E be effectively and efficiently conducted with an agile software development environment? How will future OT&E process differ from current practices? What are the changes to authorities of major stakeholders with the Naval Aviation Acquisition Enterprise? What changes to funding will be required to execute this new OT&E process?

C. WHY THIS RESEARCH IS IMPORTANT

The lack of clear direction and policy from OPTEVFOR to their OT&E squadrons may result in one of two outcomes for OT&E in an agile software development environment.

1. "Agile" will become a "buzz word" with no real meaning because no matter how fast software is developed, it will not be able to be tested in a way that meets the current requirements of the OPTEVFOR.
2. Software will be fielded without adequate OT&E, greatly increasing the risk of a weapon system's inability to prevail in combat due to mission ineffectiveness or failures of maintainability, reliability, and/or availability.

D. SCOPE

One of the reasons for the AAF and agile software development is the fact that there is no one size that fits all approaches that will fulfill every requirement of every program. As such this analysis will focus on the specifics of the F/A-18E/F Super Hornet and the software development of their H-16 and H-18 weapon system software by Boeing as well the OT&E of this software at Air Test and Evaluation Squadron (AIRTEVRON) NINE (VX-9). While the specifics may differ, the processes recommended in this research can be applied to other Naval Aviation programs going through agile software development such as the Delta System Software Configuration (DSSC) 4 and DSSC 5 for the E-2D at OT&E at VX-1.

E. METHODOLOGY

A literature review examines current GAO reports on testing software intensive programs. We will show while this research is correct in theory, it is too generalized to be implemented in the current requirements of OT&E. The authors conduct a case study of delivering new software to existing systems to show this process is too lengthy to be effective in an agile environment. We use a root cause analysis to show where this process is too show and how it can be changed. The authors will use their combined years of practical experience as operational testers at VX-1 and VX-9 respectively to show how the historical timeline norms of meeting the requirements and awaiting approvals can be "leaned" (made more efficient and effective by deleting process waste) to streamline the timeline while ensuring adequacy of test to ensure only systems that are suitable, sustainable, and available make it to frontline warfighters.

In order to answer the question about how to make OT quicker and more streamlined, while not losing sight of the reasons for OT, Chapter II of this thesis goes into background information about the OT&E process and how it currently stands. Chapter III is a literature review about how the current acquisition process works for systems that are software intensive; the literature reviewed is more much broad and does not go into testing considerations.

Chapter IV is our analysis of how OT&E could and should be done and what changes would need to be made in order to make these solutions viable. This plan has already been briefed to various stakeholders with positive feedback. Finally, Chapter V includes our conclusions and recommendations as well as some areas which warrant further study and analysis.

II. BACKGROUND

In order to understand the recommendations on how OT&E can become more agile, it is important to understand the OT&E process where it stands currently. Chapter II covers some basic background information about the current OT&E process, how long those process take and why that would be insufficient in an agile software development environment.

A. MISSION-BASED TEST DESIGN

OPTEVFOR uses a mission-based test design (MBTD), as shown in Figure 1, to create an appropriate evaluation strategy for the system under test (SUT). This will start with the Navy required operational capability/projected operational environment (ROC/POE) mission areas and then examines "the specific mission contributions ascribed to the system" (Operational Test and Evaluation Force, 2020).

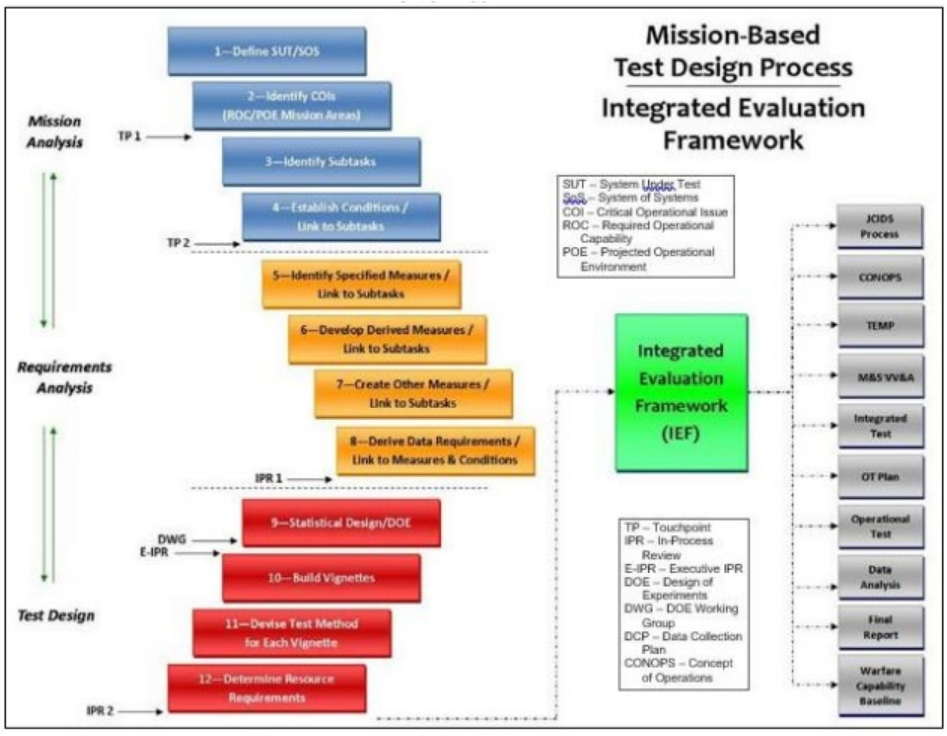


Figure 1. Mission-based Test Design Process. Source: Operational Test and Evaluation Force (2020).

As shown in Figure 1, the MBTD is divided into three phases, further divided into 12 steps. The output goal of the MBTD is to write the Integrated Evaluation Framework (IEF). These steps are sequential, but multiple steps can be worked in parallel (OT&E Manual, 2020). The first phase (steps 1–4) are about defining what needs to be tested. This is done by defining the system under test (SUT) as well as the critical operational issues (COIs) that need to be tested. It is important to clearly define these areas so OT&E can focus their efforts clearly and avoid spending time or other resources testing something that has already passed OT&E and or is not yet ready for it.

Phase 2 (steps 5–8) is creating the measures of effectiveness/suitability and data requirements to address those measures. These steps will define what needs to happen with the system in order to call the test successful. It establishes clear benchmarks that test results can be judged objectively by anyone reviewing the data.

Phase 3 (steps 9–12) are about the specific ways to collect data to address those measures. Questions addressed include the following:

- What types of environments will the test be conducted in?
- What will the scenarios (or vignettes) look like?
- What data will be collected during each vignette, etc.?

Depending on the scale of the acquisition program and the associated SUT, an appropriate pathway or multiple pathways will be used within the AAF. Historically, the acquisition of new weapon systems for the F/A-18E/F's fall into the Major Capability Acquisition Pathway as shown in Figure 2.

Figure 2. Major Capability Acquisition Pathway. Source: Under Secretary of Defense for Acquisition and Sustainment (2020a).

B. MAJOR CAPABILITY ACQUISITION PATHWAY

The majority Major Defense Acquisition Programs (MDAPs) are assigned an Acquisition Category (ACAT) based on the amount of money projected to be spent. ACAT 1 programs require oversight and T&E program approvals beyond OPTEVFOR

including congressional oversight via the Director Operational Test and Evaluation (DOT&E). For more information on ACAT programs, see Figure 3.

ACAT		

ACAT I	• MDAP[1] (Section 2430 of Title 10, U.S.C.)
	  ○ Dollar value for all increments of the program: estimated by the DAE to require an eventual total expenditure for research, development, and test and evaluation of more than $525 million in Fiscal Year (FY) 2020 constant dollars or, for procurement, of more than $3.065 billion in FY 2020 constant dollars
	  ○ MDA designation
	• MDA designation as special interest[3]
		ACAT ID: DAE
		ACAT IB: SAE[2]
		ACAT IC: Head of the DoD Component or, if delegated, the CAE

ACAT II	• Does not meet criteria for ACAT I
	• Major system (Section 2302d of Title 10, U.S.C.)
	  ○ Dollar value: estimated by the DoD Component head to require an eventual total expenditure for research, development, and test and evaluation of more than $200 million in FY 2020 constant dollars, or for procurement of more than $920 million in FY 2020 constant dollars
	  ○ MDA designation (Section 2302 of Title 10, U.S.C.)
		CAE or the individual designated by the CAE[4]

ACAT III	• Does not meet dollar value thresholds for ACAT II or above • Is not designated a "major system" by the MDA	Designated by the CAE[4]

1. Unless designated an MDAP by the Secretary of Defense (SecDef), AIS programs[5], Defense Business System programs, and programs or projects carried out using rapid prototyping or fielding procedures pursuant to Section 804 of Public Law (PL) 114-92, do not meet the definition of an MDAP.
2. ACAT IB decision authority is assigned pursuant to Section 2430 of Title 10, U.S.C. Paragraph 3A.2.b. provides DoD implementation details.
3. The Special Interest designation is typically based on one or more of the following factors: technological complexity; congressional interest; a large commitment of resources; or the program is critical to the achievement of a capability or set of capabilities, part of a system of systems, or a joint program. Programs that already meet the MDAP thresholds cannot be designated as Special Interest.
4. As delegated by the SecDef or Secretary of the Military Department.

Footnotes

5. An AIS is a system of computer hardware, computer software, data or telecommunications that performs functions such as collecting, processing, storing, transmitting, and displaying information. Excluded are computer resources, both hardware and software, that are: embedded as an integral part of a weapon or weapon system; used for highly sensitive classified programs (as determined by the SecDef) or other highly sensitive information technology programs (as determined by the DoD Chief Information Officer; or determined by the DAE or designee to be better overseen as a non-AIS program (e.g., a program with a low ratio of research, development, testing, and evaluation funding to total program acquisition costs or that requires significant hardware development). An AIS that breaches the dollar thresholds in Section 2302d of Title 10, U.S.C., as adjusted, is a "major system."

Figure 3. ACAT Programs. Source: Defense Acquisition Visibility Environment (2022).

MDAPs using the Major Capability Acquisition Pathway follow a similar approach designed to support:

> Acquisition and product support processes, reviews, and documentation will be tailored based on program size, complexity, risk, urgency, and other factors. Software-intensive components may be acquired via the software acquisition pathway, with the outputs and dependencies integrated with the overall major capability pathway. (Under Secretary of Defense for Acquisition and Sustainment, 2020b)

Regardless of the specifics of the program, there are milestones that have certain requirements that must be completed before the program is given the approval to move

on by the milestone decision authority (MDA). Part of these requirements are T&E management documents that must be reviewed, edited, and approved by both OPTEVFOR and DOT&E ACAT 1 programs.

C. AAF IMPLEMENTATION

To achieve a "decisive and sustained U.S. military advantage" through the delivery of weapon systems via the Defense Acquisition System (DAS) the DoD is directed via instruction 5000.2 to employ the AAF (Under Secretary of Defense for Acquisition and Sustainment, 2020a). As the name implies, the AAF is designed to have multiple "pathways" that acquisition programs can follow to most quickly and efficiently field a new technology based on the timeline for delivery. Within the AAF software acquisition has its own pathway, as shown in Figure 4.

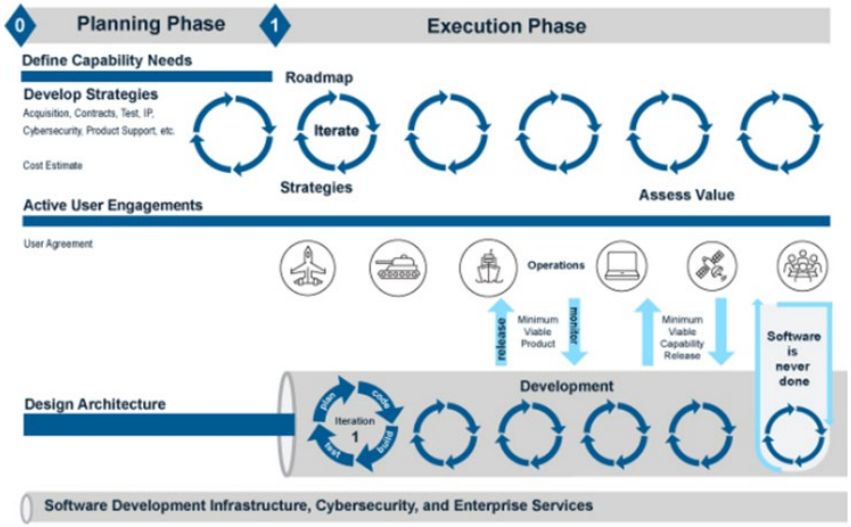


Figure 4. The Software Acquisition Pathway. Source: Under Secretary of Defense for Acquisition and Sustainment (2020b).

Note that there are two program types within the SWP. The first type is for application software which includes commercial hardware (including modified hardware) and cloud computing platforms. The second distinction is for embedded software. The second pathway provides:

> Rapid development, deployment, and insertion of upgrades and improvements to software embedded in weapon systems and other military-unique hardware systems. The system in which the software is

embedded could be acquired via other acquisition pathways (e.g., major capability acquisitions).

While the Software Acquisition Pathway (SWP) is useful for providing guidance on the implementation of software, software is rarely fielded in a vacuum. Software on major weapons systems like the F/A-18, will always be part of other acquisition programs. These instructions are useful for guidance on the general implementation of the AAF into acquisition; however, there is little specific guidance on testing, especially OT&E, for an acquisition program within the AAF that relies primarily on the SWP.

1. Best Practices Memo #32

In July of 2022, OPTEVFOR released Best Practices #32 relating specifically to the SWP. While the instruction is supportive of SWP within the AAF, it admits that programs that "utilize the SWP authority as part of a larger acquisition strategy" are still subject to all the same oversight and those requirements have not changed (Operational Test and Evaluation Force, 2022). In fact, it goes on to say that the MBTD "does not fundamentally change for SWP programs, however it must be tailored to *reduce the number of formal meetings* so that test design is completed more rapidly" (Operational Test and Evaluation Force, 2022). Figure 5 from the same document shows that even a tailored SWP still does not reduce many formal meetings for milestones or touch points.

SWP Mtg	Tailored Process Steps	Participants								
		OTD/CSH	LTE	CTF	01C AO	01D rep	ACOS*	01B*	01C*	01D*
1	TP-0	X	X[1]	X	X	X				
2	TP-1 / TP -2 / CP-0	X	X[1]	X	X	X				
3	IPR-1 / CP-1	X	X	X	X	X	X[1]	X	X	X
4	DWG/IPR-2 (if required)	X	X	X	X	X	X[1]	X	X	X
5	TP-A/C/D/E / CP-2	X	X[1]	X	X	X				
6	TP-B / TPRB / CP-3	X	X	X	X	X	X[1]	X	X	X
7	Scoring Board	X	X[1]	X	X	X				
8	CEWG	X	X[1]		X	X				
9	SERB[2]	X	X		X	X	X[1]		X	X

X – Required X – As needed X – Optional * or designated representative

[1] Mtg decision maker

[2] SERB (and ESERB, if applicable) are only applicable for QRAs and IOTE/FOTE

Figure 5. Tailored List of OPTEVFOR Milestones/Touch Points/Cyber Check Points. Source: Operational Test and Evaluation Force (2022).

Again, these tailored requirements hardly cut down on "formal meetings" and that is just with OPTEVFOR. It gives no example or suggestion about where to make cuts to all the formal meetings with stakeholders outside OPTEVFOR for acquisition programs that require the additional oversight provided by DOT&E or other agencies.

2. T&E Oversight

a. Operational Test and Evaluation Force (OPTEVFOR)

OPTEVFOR is the Operational Test Authority (OTA) for the Department of the Navy and is responsible for providing "independent and objective evaluation of the operational effectiveness and suitability of naval aviation systems" (OPTEVFOR, n.d.).

b. Director, Operational Test and Evaluation

ACAT 1 acquisition programs require congressional oversight and DOT&E provides T&E oversight for all "covered programs." DOT&E is responsible for:

> issuing DoD OT&E policy and procedures; reviewing and analyzing the results of OT&E conducted for each major DoD acquisition program; providing independent assessments to SecDef, the Under Secretary of Defense for Acquisition and Sustainment (USD(A&S)), and Congress;

making budgetary and financial recommendations to the SecDef regarding OT&E; and overseeing major DoD acquisition programs to ensure OT&E is adequate to confirm operational effectiveness and suitability of the defense system in combat use. (The Office of the Director of Operational Test and Evaluation, n.d.)

3. Required Documents

a. Integrated Evaluation Framework

Per the OT&E manual the purpose of the IEF is to:

> This product provides the OT measures, test vignettes, resources and other material required for TEMP inputs and OT Plan development. Though the IEF is an OPTEVFOR-approved document, all participating stakeholder inputs are formally resolved prior to COMOPTEVFOR approval. Disagreements between stakeholders that continue beyond the most important review meetings are adjudicated through the Running Comment Resolution Matrix (RCRM). (Operational Test and Evaluation Force, 2020)

Six reviews are done at OPTEVFOR as the IEF is being produced to ensure stakeholder alignment. These reviews are working level with O–6 level representation from the appropriate warfare division at OPTEVFOR. Members of the "core team" in these reviews include the aviation squadron's operational test director (OTD), lead test engineer (LTE), and subject matter experts (SME) to ensure an operational representative environment (Operational Test and Evaluation Force, 2020). The IEF is OPTEVFOR's biggest contribution to the Test and Evaluation Master Plan (TEMP).

b. TEST AND EVALUATION MASTER PLAN

The TEMP is "the single most important T&E document associated with an acquisition program; the controlling T&E management document" (Operational Test and Evaluation Force, 2020). In addition, the TEMP is directive in nature and has been reviewed, concurred with, or approved by all major stakeholders. It contains agreed to solutions for the cost, schedule, and performance trade space in order to conduct OT&E. The primary purposes of the TEMP include:

- Combines the Developing Agency's (DA) DT&E strategy and COMOPTEVFOR's OT&E strategy into one integrated master strategy. Because the PEO/DA and COMOPTEVFOR have independent authority,

within their respective areas, to determine program test periods and test resources, it is imperative that these independent efforts be integrated.

- Formal commitment among all stakeholders for the test approach for the life of the program. Any differences between the DA and COMOPTEVFOR on the objectives, timeline, or resources for testing have been satisfactorily resolved.
- Direction to conduct the specified T&E program, including the sponsor's committed support, and approval of the COIs.
- Provides DON T&E Executive (N94) concurrence (ACAT I through III TEMPs) on the following:
    - The thresholds and objectives as stated in the TEMP Part I are consistent with CNO approved requirements.
    - The scope of testing makes appropriate use of the Research, Development, Test, and Evaluation (RDT&E) funding, which CNO must provide.
    - The planned commitment of Fleet units for testing is consistent with CNO directed schedules and priorities. (Operational Test and Evaluation Force, 2020).

c. Test Plans

Test plans required by OPTEVFOR adds specifics that are not included in either the IEF or the TEMP. These include "dates and location of the test, test assets and ranges, squadron number, aircraft type(s), ship name/hull number, support asset type and unit name/number, detailed scenarios, etc." (Operational Test and Evaluation Force, 2020).

Test plans will also include test limitations. While the goal is always to create an environment as close as possible to the intended operational environment, this is not always possible due to many factors including asset availability, funding, range availability, or weather. These limitations and their impact on the SUT will be described in the test plan.

4. Traditional Timelines

The amount of time it takes to get approval for these required documents routed through all agencies that have oversight on OT&E is one of the largest barriers to making OT&E more agile. In the authors' practical experience getting IEFs and TEMP approved in less than six months is the exception and not the rule. The larger the program, the longer the routing process takes. Test plans that require DOT&E approval require a brief

no later than 180 days prior to the start of test and must be approved by OPTEVFOR first (OPTEVFOR, 2020). This means that it takes closer to a full year to get a test plan approved. Any timelines in writing seem to be overly optimistic. See Table 1 for more information on timelines of signatures for OPTEVFOR only.

Table 1. Signature Authority Timelines. Source: OPTEVFOR (2020).

Table 3-2. Signature Authority
UNCLASSIFIED

T&E Document	Response Time	Brief Required	Signature Authority	
			N00	DIV Director
TEMP and T&E Strategy	15 working days (Note 1)	No (Note 1)	X	
Oversight test plans (Note 2) (Includes IOT&E, FOT&E, OA, EOA, and Multiservice Operational Test and Evaluation (MOT&E) oversight test plans)	60 days prior to test	COT Brief only	X	
All evaluation report letters (Includes MOT&E Test Reports). Note that TD signs Data Analysis Summary. Div. Director signs Deficiency cover letter.	60-90 days after test (Note 3)	No. Covered by ESERB	X	
Interim Reports	As required	Yes	X	
VCD messages/reports in which COI resolutions are being changed from the previous phase of test	35 days after test	No. Covered by ESERB	X	
Quick Reaction Assessment (QRA) messages/reports	60 days after test	No. Covered by ESERB	X	
All OT&E support letters (Warfare Division responsible for drafting)	30 days prior to test	No	X	

Table 3-2. Signature Authority UNCLASSIFIED				
T&E Document	**Response Time**	**Brief Required**	**Signature Authority**	
			N00	**DIV Director**
Deficiency report messages	As directed	Yes	X	
M&S Accreditation Plan	ASAP after need identified in E-IPR, NLT 1 year prior to test	Yes	X	
All M&S Accreditation Letters	NLT 90 days prior to test	Yes (for programs on oversight list)	X	
IEF/IEF Revision	(Note 4)	No. Covered by E-IPR	X	
Tailored IEF where 1) Mid-Tier Rapid Fielding will be decided, or 2) RCRM requires flag-level resolution, or 3) Warfare Division Director believes flag signature is most appropriate		No. Covered by E-IPR	X	
Tailored IEF that does not meet any conditions requiring 00 signature (see above)		No		X
IA/Interoperability Assessment Reports	NLT 90 days post- test	Yes	X	
Integrated Assessment Plan (IAP)	60 days after program initiation	Yes	X	
Operational Utility Assessment (OUA), Military Utility Assessment (MUA), and Limited Military Utility Assessment (LMUA) reports	60 days after demonstration unless specified otherwise	Yes	X	
VCD messages/reports in which COI resolutions are NOT being changed from the previous phase of test	35 days after test	No		X
Level of Test Determination (LTD) Report		Yes (for programs on oversight list)		X (Note 13)
Administrative Updates to Previously Approved TEMPs	As required	No		X
Capabilities Documents, Initial Capabilities Document (ICD)/CDD/CPD Clarification Letter	As required	(Note 5)		X
TEMP comment letters (for O-6 level reviews)	30 days from receipt	Yes (Note 6)		X
O-6 level reviews of MOT&E Test Plans and Test Reports	14 days from receipt	Yes		X
Non-oversight test plans (Note 2) (includes IOT&E, FOT&E, OA, EOA, and MOT&E non-oversight test plans)	30 days prior to test	(Note 7)		X
Oversight and non-oversight QRA and VCD test plans, and IT data collection plans (Note 8)	30 days prior to test	Yes		X
Risk/Deficiency forwarding letter	Prior to the SERB	No		X
Joint Capabilities Technology Demonstration (JCTD) Demonstration Execution Document (DED)	30 days prior to demonstration	(Note 7)		X
Anomaly report messages		(Note 9)		X
TEMP input letters	90 days after program initiation	No (for programs on oversight list)		X (Note 10)
Standard/Combined DT/OT Memorandums of	30 days prior to test (at	No		X

Table 3-2. Signature Authority UNCLASSIFIED				
T&E Document	**Response Time**	**Brief Required**	**Signature Authority**	
			N00	**DIV Director**
Agreement (MOA)	test plan signing)			
IEF Change Letter	(Note 4)			X
Support documentation (Integrated Logistic Support Plan (ILSP), Navy Training Plan (NTP), etc.)	15 days from receipt	No (Note 8)		X
M&S Operational Requirement Input Letter	During IEF development, as soon as need is identified	No		X
Letters of Instruction (LOI)	30 days prior to test	No		X (Note 11)
Adjunct tester forms	30 days prior to test	No		X
DT assist MOA-(if used)	30 days prior to test	No		Division Director/ VX CO
IT MOAs and Charters	As required	No		Division Director/ VX CO
AOC letters and DT Assist Letter of Observation (LOO)	30 days after test/ demonstration	As required		X
OT commencement messages or e-mails		No		X
OT completion messages or e-mails	End of test as determined by division director	No		X
ACAT IVM & Abbreviated Acquisition Program (AAP) concurrence letters				X
OPTEVFOR Tactics Guides (OTG)	120 days after evaluation report	As required		VX CO (Note 12)

Notes:
1. Assumes a formal O-6 TEMP review has been completed and that all critical OPTEVFOR comments were satisfactorily resolved. If not, a brief to the Commander is required.
2. Commander signs all ACAT I, DOT&E oversight, and controversial test plans. Additionally, the Commander may sign all standard test plans, *when desired,* 30 days prior to testing.
3. Ninety days for ACAT I/IA and MOT&E; 60 days for all others.
4. For new programs, coordinate IEF completion to support initial TEMP development (MS-B). For existing programs, IEF must be approved in time to support next phase of test or MS. IEFs for programs on oversight list are forwarded to the DOT&E to support TEMP approval.
5. Briefs are on a case-by-case basis. The Commander may elect to sign comment letters with contentious issues.
6. Division Director shall brief Commander or Deputy on all TEMPS with critical OPTEVFOR comments.
7. Division director signs (provides a copy to Commander/Deputy for review; briefs on a case-by-case basis) standard ACAT II, III, and IVT test plans. Staff through 01A/C prior to division director signature.
8. QRA test plans for programs on oversight list are forwarded to the DOT&E. For the case of DOT&E oversight, the Commander will sign the QRA test plan.
9. Brief the Commander (or Deputy in his absence) prior to release.
10. Sign "By Direction."
11. LOIs prepared at VX/VMX/HMX may be released by the squadron Commanding Officer (CO).
12. VX COs authorized to sign "By direction." The Commander will sign controversial and special interest OTGs and all Naval Warfare Publications (NWP). Briefing requirements will be determined on a case-by-case basis.
13. If the LTD RCRM required Flag/SES-level intervention, the LTD will be signed by the Commander.

It is also important to understand that the timelines listed in the table are only for formal reviews. There is much more back and forth between the test squadrons and OPTEVFOR. As a real-life example, we looked at the test plan writing and approval

process for the E-2D DSSC 4.0 OT test plan. While the formal data was not available for the F/A-18 timelines, based on practical experience, they are very similar.

The first DSSC 4 test plan document was created on March 3, 2021 (Figure 6). This does not account for the working groups that took place beforehand to map out and strategies about what would be in the test plan. It did not enter formal routing at OPTEVFOR until August 26th, 2022 (Figure 7) and was not signed until October 25th, 2022 (Figures 8). This process took nearly two years to complete, nowhere near what would be required in an agile software development environment.

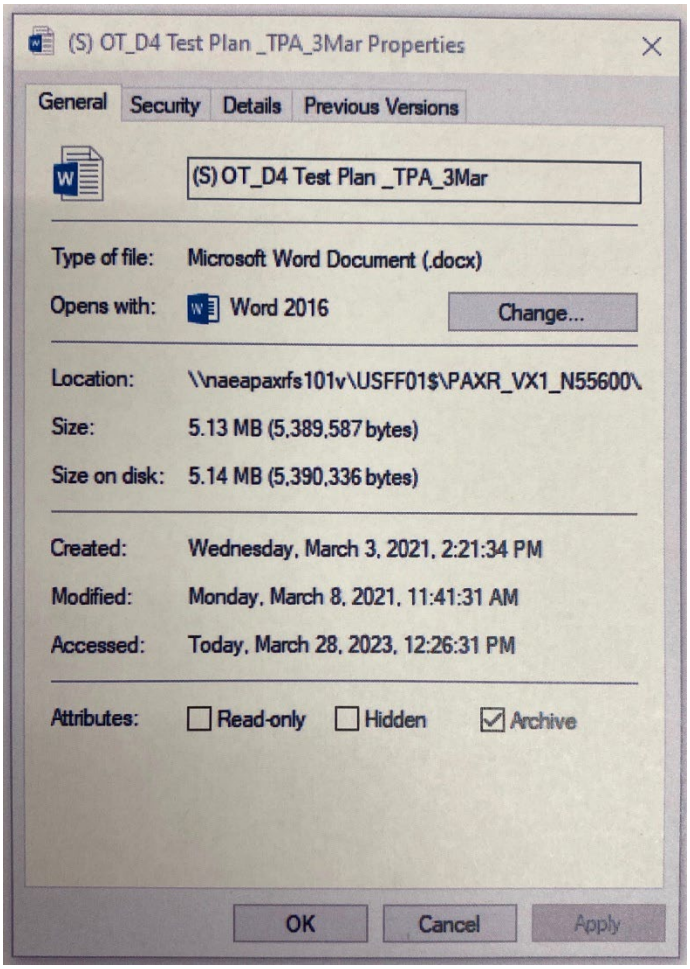
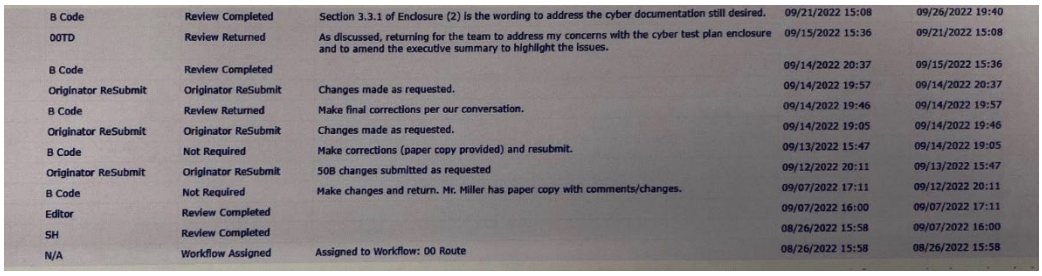
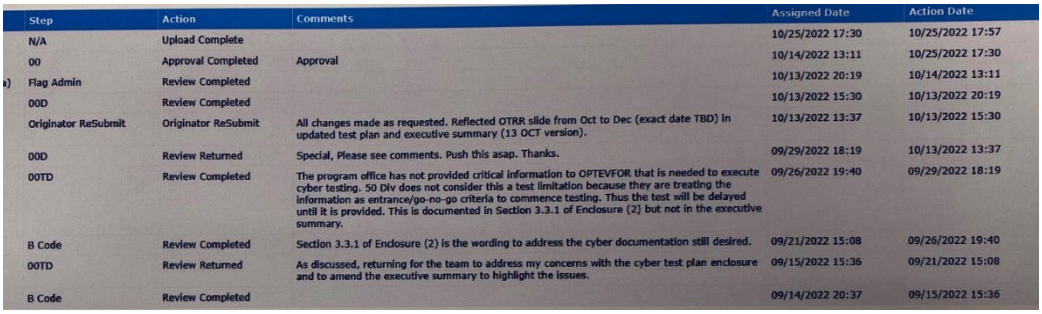


Figure 6. OT DSSC 4 Document Creation



B Code	Review Completed	Section 3.3.1 of Enclosure (2) is the wording to address the cyber documentation still desired.	09/21/2022 15:08	09/26/2022 19:40
00TD	Review Returned	As discussed, returning for the team to address my concerns with the cyber test plan enclosure and to amend the executive summary to highlight the issues.	09/15/2022 15:36	09/21/2022 15:08
B Code	Review Completed		09/14/2022 20:37	09/15/2022 15:36
Originator ReSubmit	Originator ReSubmit	Changes made as requested.	09/14/2022 19:57	09/14/2022 20:37
B Code	Review Returned	Make final corrections per our conversation.	09/14/2022 19:46	09/14/2022 19:57
Originator ReSubmit	Originator ReSubmit	Changes made as requested.	09/14/2022 19:05	09/14/2022 19:46
B Code	Not Required	Make corrections (paper copy provided) and resubmit.	09/13/2022 15:47	09/14/2022 19:05
Originator ReSubmit	Originator ReSubmit	50B changes submitted as requested	09/12/2022 20:11	09/13/2022 15:47
B Code	Not Required	Make changes and return. Mr. Miller has paper copy with comments/changes.	09/07/2022 17:11	09/12/2022 20:11
Editor	Review Completed		09/07/2022 16:00	09/07/2022 17:11
SH	Review Completed		08/26/2022 15:58	09/07/2022 16:00
N/A	Workflow Assigned	Assigned to Workflow: 00 Route	08/26/2022 15:58	08/26/2022 15:58

Figure 7. OT DSSC 4 Routing Page 1



Step	Action	Comments	Assigned Date	Action Date
N/A	Upload Complete		10/25/2022 17:30	10/25/2022 17:57
00	Approval Completed	Approval	10/14/2022 13:11	10/25/2022 17:30
Flag Admin	Review Completed		10/13/2022 20:19	10/14/2022 13:11
00D	Review Completed		10/13/2022 15:30	10/13/2022 20:19
Originator ReSubmit	Originator ReSubmit	All changes made as requested. Reflected OTRR slide from Oct to Dec (exact date TBD) in updated test plan and executive summary (13 OCT version).	10/13/2022 13:37	10/13/2022 15:30
00D	Review Returned	Special, Please see comments. Push this asap. Thanks.	09/29/2022 18:19	10/13/2022 13:37
00TD	Review Completed	The program office has not provided critical information to OPTEVFOR that is needed to execute cyber testing. 50 Div does not consider this a test limitation because they are treating the information as entrance/go-no-go criteria to commence testing. Thus the test will be delayed until it is provided. This is documented in Section 3.3.1 of Enclosure (2) but not in the executive summary.	09/26/2022 19:40	09/29/2022 18:19
B Code	Review Completed	Section 3.3.1 of Enclosure (2) is the wording to address the cyber documentation still desired.	09/21/2022 15:08	09/26/2022 19:40
00TD	Review Returned	As discussed, returning for the team to address my concerns with the cyber test plan enclosure and to amend the executive summary to highlight the issues.	09/15/2022 15:36	09/21/2022 15:08
B Code	Review Completed		09/14/2022 20:37	09/15/2022 15:36

Figure 8. OT DSSC 4 Routing Page 2

Note: While the title of the document is portioned marked as “Secret” in the figures, that is in reference to contents of the document, not the name of the document or the timeline of its routing.

5. Agile Development and the Scaled Agile Framework

DoDI 5000.87 (Operation of the Software Acquisition Pathway) provides policy guidance on the development and sustained procurement of embedded software systems. By utilizing modern-iterative software development processes, end users should be provided with incremental capability updates at a faster rate. Agile has been the term most associated with this reform in software development, but what is agile development?

Agile software development at its heart is a simple process philosophy developed into a four-line manifesto, it values the rapid development of usable software that provides end-user value over the traditional planning methods and processes in waterfall development. This allows for software to easily adapt to changes in the market throughout development. By pushing for more collaboration across stakeholder groups

during development, software would be deployed more often, with better value to end-users.

To implement this change to modern-iterative software development processes, PMA-265, the advanced weapons lab (AWL), and Boeing have chosen to implement the scaled agile framework (SAFe) for Government enterprises (Brey, 2021). The SAFe framework allows large, established organizations to quickly transition from traditional waterfall development methods to lean-agile processes by realigning personnel into a team-of-teams organization governed by a standard, short development timelines. This allows for each team to quickly plan and self-organize to tackle required development with an understood end date. This reorganization at Boeing, the AWL, and PMA-265 has allowed, from a developmental standpoint, an ability to release smaller software configuration set (SCS) updates every six months.

D. ADDITIONAL STAKEHOLDERS

Anyone with a vested interest in the test outcome or responsibility for assuring that the testing completed was adequate can be considered a stakeholder. This includes businesses and private companies responsible for delivery of a system all the way to the warfighter who will ultimately be the end user. While the stakeholders will differ for each project, the following are stakeholders specific to the delivery of the F/A-18's SCS.

1. OFFICE OF NAVAL OPERATIONS-AIR WARFARE DIVISION (OPNAV N98)

N98 is the resource sponsor for Naval Aviation, including the F/A-18 Super Hornet. They are heavily involved in Joint Capabilities Integration and Development Systems (JCIDS). They ensure there is "building, integrating, and defending yearly Program Objective Memorandums (POMs) for all Naval Aviation Programs" (Blickstein et al., 2016). N98 is kept informed about the F/A-18's priorities to ensure appropriate funding is in place for the acquisition of new systems. The Requirements Officers (RO) work directly for N98.

2. Projects (Program) Manager, Air (PMA)

PMA-265 is the primary program office and is responsible for acquiring, delivering, and sustaining the F/A-18 and EA-18G aircraft (*PMA-265,* n.d.). It accomplishes this mission by utilizing the funding and requirements developed by the ROs at OPNAV 98 and contracting with various organizations for the bulk of software development. In the case of the F/A-18, Boeing is the prime contractor. Additionally, PMA-265 partners with government testing and development resources at the Advanced Weapons Laboratory to ensure proper integration with other programs of record. The PMA manages the acquisition program baselines to ensure the funded capability reaches the Fleet in-line with the operational need date at the speed of relevance.

3. Original Equipment Manufacturer (OEM) Contractor (Boeing)

Boeing is the prime contractor of the F/A-18 and EA-18G and the primary developer and integrator of the SCS software. Boeing works to contract and develop the requirements from PMA-265 and other PMAs that must integrate onto the greater airframe and baseline SCS. Additionally, they work in conjunction with government software integrators at Naval Air Warfare Center-Weapons Division (NAWCWD).

4. Naval Air Warfare Center Weapons Division (NAWCWD)

NAWCWD is the government integrator and developer that has access to "state-of-the-art and one-of-a-kind laboratories and facilities that enables the delivery of unparalleled weapons research, development, acquisition, test, and evaluation." NAWCWD is the prime government integrator for the F/A-18 and EA-18G and one of the main service providers to PMA-265. The engineering and test teams also work in close concert with Boeing to ensure delivery of updated software for PMA-265.

5. Developmental Test (DT) Squadrons

Responsible for testing to ensure contract fulfillment and compliance to contract specifications. Their testing will be done in controlled environments working closely with the OEM and other contractors. The F/A-18 DT squadrons are VX-23 and VX-31.

6. Operational Test Squadrons

Responsible for testing to ensure operational effectiveness and suitability to include requirements of availability, maintainability, and reliability are met. Their testing will be done in operational realistic environments by "fleet average" users. Figure 9 breaks down into more detail OT vs. DT squadrons.

	DT&E	OT&E
Managed by	Developer (contractor or government PM)	Independent government test agency
Primary objectives	• Test to specification requirements • Precise performance against threshold measurements	• Test to operational requirements • Performance measurement of operational effectiveness and suitability
Testing assets	Developmental test articles or prototypes	Production representative systems
Conducted by	Trained, experienced operators aided by technicians with contractor involvement	Typical users with normal training and restricted contractor involvement
Environment	Highly controlled	Realistic/combat conditions with operational scenarios

Figure 9. Developmental Test and Evaluation to OT&E Comparison. Source: Mortlock et al. (2009).

This background chapter provides context to the reader to understand the vastness of the Navy OT&E organization. This enterprise extends and is interlocked with many other government agencies including the other service branches that make this process all the more complicated. It would be difficult for someone not fully involved in this process to come in and change it, because they wouldn't understand the complexities of it all.

The next chapter reviews what has been written on acquisition of software intensive programs and how OT&E fits within the Defense acquisition process.

III. LITERATURE REVIEW

Since the introduction of the AAF and the iterative and continuous software development cycle, there has been lots of research on how to implement these ideas into software intensive programs. There is also a lot of research done before the AAF stating that software intensive programs were difficult to test and implement. While these studies show the difficulty in testing software intensive programs, and the need to test early and often, they are too general and do not address how OT&E can be streamlined to maintain pace with the speed of relevance.

A. GAO REPORTS

In 2014, the GAO reported that the F-35 would most likely be delayed due to delays in software development and software delivery with less than expected capability as shown through DT&E. At the time DOT&E estimated the program was at least 13 months behind schedule, and that the total cost of the program was over $1 trillion, which DoD officials determined to be "unaffordable" (GAO, 2014).

This report goes on to say how different it is to test software intensive programs due to the complexity of how software affects everything in a system. Software in the F-35 program would often be delayed and then delivered with limited capability. Software would then need to be retested several times because when new capability was delivered it would affect other systems that had already passed DT&E. This retesting was expensive and time-consuming.

In order to avoid the need for retesting, another GAO report from 2019 suggests that including users earlier in the process of both incremental and iterative software development could benefit software intensive programs. While this is certainly true, the report's suggestions are too general to fit into OT&E requirements as they stand today. It suggests that "testing" and "user feedback" are just two steps in the process that should take no more than one year, while not acknowledging that those steps alone would take longer than one year under current requirements (GAO, 2019). The systems from this study ultimately went on to fail OT&E due a number of reasons, but officials from the Defense Digital Service, DOT&E, and other DoD leadership did acknowledge that "the

rapid development of software using newer software practices does not fit with the current requirements of the DoD acquisition process" (GAO, 2019).

B. CONCLUSIONS

While multiple studies have been done on how to field software intensive programs under the umbrella of the AAF, there have been no studies to address how to cut down the bureaucratic requirements of OT&E in a way that can match the timelines required. Nearly all new systems fielded can be considered "software intensive" and OT&E needs to learn their processes if they are going to be able to continue to ensure adequate testing in an agile environment.

IV. ANALYSIS FOR AGILE OT&E

This chapter will focus on the author's perspective about how to make OT&E more agile to be able to keep pace with typical software development timelines, while ensuring adequate testing is done and the fleet can have confidence in what they are being delivered.

Our root causes were primarily barriers to communication including limited personnel, and a lack of clear chains of commands when programs had pieces spread across multiple chains of command. We then show specifically how these root causes affected the H-16 software testing for the F/A-18. Figure 10 is a visual depiction of these root causes.

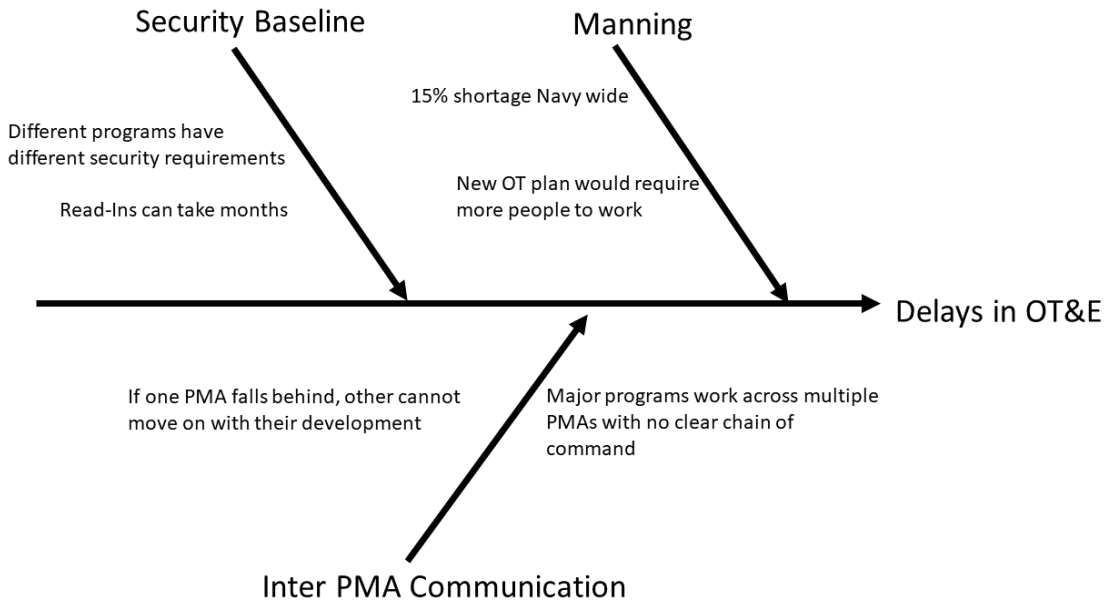


Figure 10. Fishbone Diagram

A. ROOT CAUSES

1. Appropriately Manned Squadrons

Currently both VX-9 and VX-1 are understaffed from both an aircrew and Flight Test Engineers (FTE) perspective. This Navy wide issue is not unique to the test community. In 2020, Navy ships were undermanned by 15% on average across the fleet

(GAO, 2022). Our proposal would require even more manning than the current allotment at either squadron. This would facilitate a "blue and gold" crew mentality where half the squadron could be working on testing, analysis, and reporting, while the other half of the squadron could be planning for the next cycle of upcoming test.

More manning across the squadron would decrease the time for everything. 15% more maintainers mean less downtime of the jet allowing for more sorties to gather test data. More aircrew means that the same pilots aren't flying everything single event allow for more operator input to design and analysis considerations. More engineers allow for faster analysis of in-flight data and more eyes to look at problems and coming up with faster and better solutions.

2. Appropriate T&E Security Baseline

One of the biggest hurdles to OT&E from a mission planning and data sharing perspective is making sure all parties involved have the appropriate security clearances in order to be fully involved in the process. This becomes even more of an issue in the Joint environment, with multiple Services and multiple warfare domains. In order for OT&E to be effective there must be a T&E security baseline that everyone can be read into to allow the flow of information across the DoD enterprise.

Nothing is tested in a bubble. The E-2D Hawkeye provides command and control for various platforms including the F/A-18 Super Hornet. The issue is that neither platform allows the other to be fully read-in to all their programs. This leads to numerous joint test events when each platform has their own test objectives, they can't tell the other about. When both platforms have different secret objectives they are working towards it is difficult to work together during limited combined flights.

Other problems that arise from security concerns are the sharing of post flight data. When platforms are designed to work together it is almost certain that some kind of data from both platforms will be required to validate test points. Data from different platforms are stored on different IT systems that don't often talk to each other.

3. Effective Inter-PMA Coordination

PMA-265 is responsible for the F/A-18 Super Hornet and PMA-231 is responsible for the E-2D Hawkeye airframe. Both programs have systems on them that belong to other PMAs. For example, PMA-298 is responsible for Special Programs and Mission Integration is heavily involved in both the F-18 and E-2D. There are also PMAs for various weapon systems, weapons, and joint hardware.

While sometimes there is a clear line of authority with PMAs in the form of Program Executive Offices (PEOs), most of the time the PMAs are independent. In order for our proposal to work any PMA involved in test must be open and honest and provide constant communication to the other PMAs to align efforts to decrease duplication of work and increase interoperability.

This type of communication is not limited to just the PMAs, but to all stakeholders involved. If one group falls behind a timeline, that must be communicated to all other stakeholders in order for the enterprise to realign their priorities and timelines. This must not be viewed as a failure by the party responsible, but rather as just what it is, open communication to keep everyone informed and make the best decisions easier to come by.

One clear example of the roadblock that come up when systems go across multiple PMA's is with Multifunctional Information Distribution System (MIDS) run by PMA 101. Nearly every platform in the Navy uses MIDS to maintain a common operating picture (COP). In the E-2D specifically the MIDS boxes work in conjunction with the tactical aid to navigation (TACAN). This can cause a safety of flight issue when MIDS is turned off behind the boat prior to landing and the E-2D loses TACAN resulting in diminished situational awareness on aircraft own position during terminal phases of flight. It is ultimately unclear who is supposed to fix this problem.

On top of those MIDS issues all platforms that use MIDS have to deal with the Federal Aviation Administration (FAA) because MIDS works on some of the same frequencies that the FAA owns. There is no clear written chain of command between the FAA and the DoD and as a result each platform that uses MIDS is forced to work

independently with the FAA in order to radiate MIDS. This process is burdensome for all involved and extremely inefficient since no one ones all these systems.

After solutions to these root problems are in place our solutions would be much more viable and realistic. In order for stakeholders outside of PMA-265 to understand our recommendations and be able to apply them to their programs it is important to use the same terms for the same things.

B. PMA-265 SAFE DEFINITIONS

While different organizations may use the same terminology for different things, we will be using the definitions generally accepted and used by PMA-265 in the following ways.

1. Initiative

Very similar to a statement of requirement (SOR). This is the most basic unit of what a new delivery must do. A traditional SOR can be broken down into different initiatives and the same initiative can fulfill different SORs.

2. Mosaic

A collection of initiatives that create usable capability to the fleet. Mosaics are what "move the needle" for the warfighter. Mosaics will make the fighter more lethal or more survivable in some fashion. Mosaics should be associated with the closing of capability gaps that exist within a strategic theme.

3. Strategic Theme

A collection of mosaics that creates capability for a specific warfare mission. An example of a strategic theme could include long range organic weapons, air-to-air radar, or sustainment.

4. Minimum Viable Product (MVP)

The lowest level of new capabilities that increase Fleet capability. MVPs will be made up of initiatives and/or mosaics and may fall into one or more strategic theme. The goal of the MVP is to become a Minimum Viable Capability Release (MVCR).

5. MVCR

PMA-265 uses the same definition for MVCR as DoDI 5000.87 stating that:

> The initial set of features suitable to be fielded to an operational environment that provides value to the warfighter or end user in a rapid timeline. The MVCR delivers initial warfighting capabilities to enhance some mission outcomes. The MVCR is analogous to a minimum marketable product in commercial industry. (Under Secretary of Defense for Acquisition and Sustainment, 2020b)

In this research project, the "user" will be considered the Fleet, and the "customer" will be OPNAV N98.

After these definitions are understood we can move onto the processes and products we recommend in order to streamline the OT&E process in an agile software development environment.

C. PULVERIZER

The Pulverizer is a new requirements decomposition event that looks to develop and prioritize strategic themes, mosaics, and initiatives for the next two years of development. Pulverizers occur every six months and become an iterative process to ensure that the development roadmaps can adapt to urgent operational needs. The meeting was started out of necessity due to the limited capacity of Boeing and NAWCWD to develop contracted requirements. The Pulverizer strives to include all stakeholders including other PMAs, NAWCWD, Boeing, and knowledgeable end-user representatives to ensure that mosaics move the needle of capability for the Fleet.

The Pulverizer is the first step in the agile development process; every initiative, funded or not, is inserted into an appropriate strategic theme based on its capabilities. After the initiatives are binned by strategic theme, mosaics are developed that create usable end-user capability to cover the capability gaps associated within the strategic theme. Figure 11 is an example of the requirements decomposition. This process continues until all initiatives are placed into a mosaic. At the end of the Pulverizer all mosaics are then voted on for overall prioritization. The voting board consists of leads within the major stakeholders including: NAWDC, PMA-265, NAWCWD, and T&E.

This board ensures that operational priorities are weighed against acquisition realities to ensure the best use of limited development capacity.

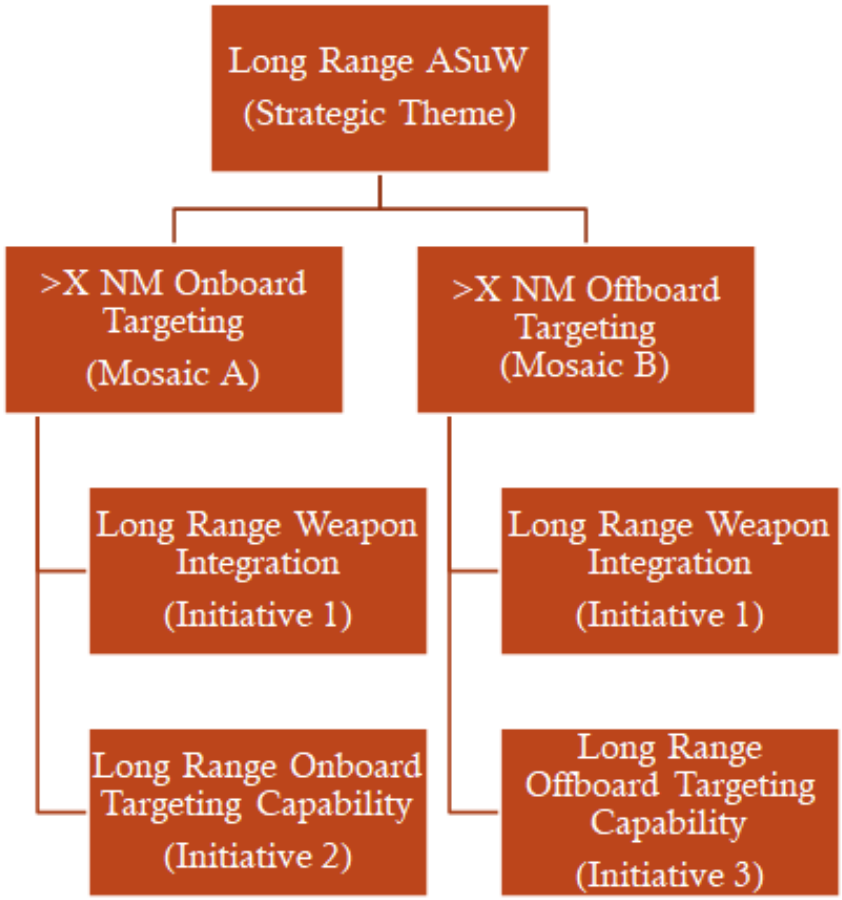


Figure 11. The Pulverizer

The main output of the Pulverizer is a prioritized list of mosaics that should be developed over the next two years. This prioritized list must be approved by not only the program office but also the appropriate requirements officers to ensure funding can be moved—one of the many problems of transitioning to the agile development process. The program office is currently striving to request the "colorless" software funding lines to limit losses that normally occur in reallocations. If the appropriate funding cannot be moved quickly to the program's priorities, this will create wasted development with limited capacity, further delaying needed operational capabilities.

D. AGILE TEMP

With the potential of moving requirements bi-annually, the traditional TEMP process will be unable to maintain the pace. Traditionally, the TEMP was an incredibly detailed document that included all the requisite SOR definitions as well as measures required to evaluate performance up to five years before test execution. With the advent of agile software development process and the Pulverizer, that detail may be missing since the SORs may have been further refined into smaller MVPs. This lack of detailed knowledge requires the Agile TEMP to be a broad stroke level-of-effort document to ensure proper funding is available when needed. This style of funding is needed to ensure that the appropriate ordnance, ranges, and targets are on hand to efficiently test as mosaics reach a mature, stable condition. This means a higher upfront cost to the program office to ensure the rapid pace of agile software testing.

E. UNIVERSAL TEST PLAN WORKING GROUP

Post Pulverizer, the development and T&E organizations should have an approved list of priority mosaics and a basic understanding of their associated development timelines. The next event will be the Universal Test Plan Working Group. The stakeholders that must be at this working group shall be: NAWCWD, DT, OT, OPTEVFOR, and DOT&E. This meeting is used to quickly develop an integrated test and evaluation strategy to maximize efficiency and to limit duplicative testing between OT and DT squadrons. Additionally, this working group should develop appropriate and agreed-to entrance and exit criteria for each phase of test to ensure an agreed upon definition of "good enough."

The output of this meeting should feed the individual test plans supporting the developmental testing, integrated operational test assessments, and dedicated operational testing. Additionally, the more detailed discussion on test strategy will allow for further refinement of the Agile TEMP.

F. INTEGRATED OPERATIONAL TEST ASSESSMENT (IOTA)

The integrated operational test assessment is a new phase and definition of integrated test. Traditionally, DT would conduct their testing and there would be an

operational test readiness review (OTRR) to determine whether or not a system was ready to go to OT. This would be DT and OT owning the process on either side of OTRR. The IOTA would blur that line to ensure both DT and OT were involved and making decisions both pre and post OT&E. In our experience IT has often over-promised and under-delivered. This is due to the differences in testing environments, equipment, and execution which prevented dual use of the collected data. This generally did not lead to efficiencies in test schedule but served more as informational on the status of the SUT between DT and OT squadrons. In the new agile software development construct, there needs to be an actual integration effort to ensure severe operational deficiencies can be quickly characterized, fixed, and re-tested to support faster Fleet releases.

An IOTA is a small, targeted, test plan/period that will look to qualitatively assess the effectiveness and suitability of the stability and capability of a ready mosaic. This will be a fully integrated effort between the DT and OT squadrons to ensure appropriately instrumented aircraft are available in operational mission-based test events to drive down deficiency characterization. This allows for the T&E enterprise to quickly find and fix potentially operationally relevant deficiencies prior to the dedicated annual OT&E period. To ensure maximum efficiency the IOTA test plan will be informed by the DT test plan as well as OPTEVFOR and DOT&E design of experiments. This should allow data obtained in the IOTA to be acceptable for the dedicated annual OT&E period.

At the completion of the IOTA period, a qualitative assessment of the critical operational issues (COIs) will be produced. This will either affirm that the mosaic is ready to be included in the annual Fleet release build or that it needs further development. The authority to make this call will be the commanding officer of the appropriate OT squadrons. This allows an informed end-user representative to agree that the development adds value to the Fleet. Additionally, the Immediate Superior in Chain of Command (ISIC) of the OT squadrons is the Commander, Naval Air Forces (CNAF) who is responsible for the manning, training, and equipping of Naval Aviation. If the mosaic has enough operational importance, the OT commanding officers can brief CNAF on whether the mosaic build should be Fleet released without the full regression testing of the software.

G. DEDICATED ANNUAL OT&E (DAOT&E)

Every year there will be an Operational Test Readiness Review (OTRR) by the last day of June. While currently the OTTR process does take time for ACAT1 programs by having it the same time every year, it should be much easier to coordinate across OPTEVFOR and DOT&E. They should also have a very good idea about how OTRR will go based on the IOTAs that have been completed previously during the year.

This creates a regular goal for all the developmental, logistical, and techniques, tactics, and procedures (TTP) development processes. All the mosaics that appear promising during the IOTA phase will go through a "build hardening" process so that if OTRR takes place on June 30, then on July 1 the software can be loaded into the jets.

Perhaps more importantly, this will ensure that fleet builds are ready to be tested every year in the Gray Flag Exercise traditionally held every year in August. Gray Flag is the premier OT&E exercise where anyone in the DoD can come and participate in one of the most realistic and operationally representative environments that can be created in the United States. Over a two- or three-week period, there will be multiple large force exercises (LFEs) per day involving all the services in a land, air, and maritime environment.

Gray Flag provides a unique opportunity to test systems in an interoperative way as well as getting into high level tactics such as denied and jamming electronic attack (EA). These environments would be nearly impossible for any one squadron or organization to replicate on their own. It is also extremely efficient from a cost perspective since it is a "pay to play" mentality and anyone that wants to participate has to fund their own way, but then they have immediate access to high level systems and assets they could never pay for on their own. A stable, predicable DAOT&E, such as Gray Flag, is essential to OT&E in an agile environment.

The output goal from DAOT&E will be formal test reports signed by OPTEVFOR, and Operational Tactics Guide (OTG) for the fleet, COI resolution, and blue/gold sheet reporting. These reports will be critical because they will provide the documentation needed by the fleet to properly train their new users on what to expect with the new software. By having the DAOT&E period be in the summer, this allows

approximately six months to execute tests, develop new tactic execution, and write formal reports for a fleet release on January 1 the following year.

This chapter covered our analysis of what from the current OT&E process would be unsustainable in an agile software environment, and what changes could be made in order to correct those deficiencies. We started by doing a root cause analysis and found the major root causes to be improperly manned squadrons, an appropriate security baseline and effective inter-PMA communication. In order to address these root causes we proposed 5 initiatives that could significantly reduce or illuminate the slowdowns faced by OT&E today. Those initiatives are the Pulverizer, to better define requirements among all stake holders, and an agile TEMP to be more high level and less specific. We also recommend the Universal Test Working Group to bring all the stake holders together on a regularly scheduled basis and the IOTA to bring down the decision-making power to lower levels to commanders who are involved in the testing in the day to day and have a much better picture of what is ready for test. Finally, we recommended the dedicated OT periods such as Gray Flag to ensure multiple platforms and programs come together to test to ensure interoperability in an operationally relevant environment. By following these recommendations, it would be much easier to get the OT&E process on a predicable yearly schedule where one or two major software systems could be fielded.

V. CONCLUSIONS AND RECOMMENDATIONS

The previous chapters answered the questions about how OT&E could be conducted efficiently in an agile software development environment. We proposed a more predicable yearly calendar with major events in order to keep the newest software being tested in realistic environments. We studied how that process would be different than what it is today, mainly less written and administrative requirements and having the OT&E testing requirements be more agile as well. We also proposed changes to the authorities on these products to reduce the timelines of approval, and how this process would be funded.

The following section provides our research questions and the answers to our research questions from this research project.

Question 1: How can OT&E be effectively and efficiently conducted with an agile software development environment? Answer: We found that by using our 5 initiatives (Pulverizer, agile TEMP, universal test plan working group, IOTA, and annual dedicated OT&E period) OT squadrons will be just as flexible as the software development is intended to be. It also gives the added benefit of having the same schedule every year so all stake holders will be able to work off the same timelines. If something is delayed that communication needs to be clear so stake holders can test what is ready while delayed programs slide to the next on ramp.

Question 2: How will future OT&E process differ from current practices? Answer: Testing processes will become much more high-level and less specific. The documents that are required now to be approved by OPTEVFOR and DOT&E will be broad brush will the test specifics will be worked at the squadron level via the development of the mosaics and the Pulverizer process.

Question 3: What are the changes to authorities of major stakeholders with the Naval Aviation Acquisition Enterprise? The IOTA will bring the decision level of what is ready for OT down the squadron COs who have a much better understanding of program risks, timelines, and deficiencies. Since the testing will be much more regular and scheduled it won't be necessary for high level flag officers to sign off on all the risk.

Question 4: What changes to funding will be required to execute this new OT&E process? Answer: This type of testing will require new, creative ways of thinking about how money is budgeted. “Colorless” money is our recommendation since it will give more flexibility to software development, while ensuring that money for a program isn’t lost due to not being able to spend it at a specific time.

While the recommendations in the previous chapter were only the recommendations and opinions of the authors, it is important to note that all the major stakeholders were briefed on these recommendations, and they were widely met with positive feedback and concurrence. These stakeholders include OPTEVFOR, DOT&E, PMA-265, Boeing, VX-31, NAWCWD, and VX-1. The only major disagreement was who would have the authority to sign off on IOTAs. OPTEVFOR wanted it to rest with them, as did the PMA. Figure 12 is a visual depiction of how the proposed OT&E would work.

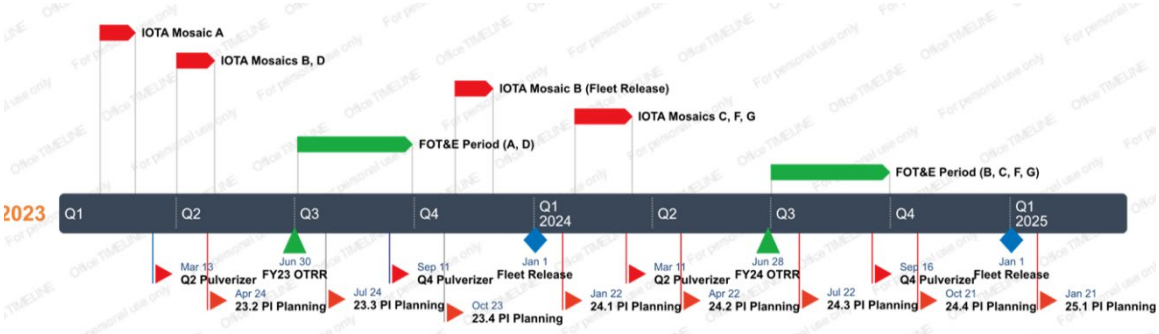


Figure 12. Proposed T&E Timeline

It is also important to note that while this is a good plan for the future, the infrastructure from the contractor (Boeing specifically in this case) is not yet in place. What is lacking is the appropriate product baseline engineering. This includes the ability to create a stable baseline where all other features are added. It also includes a “feature toggle” ability which will allow code to be developed and “in” the system without being turned “on.” This means that code can be developed, installed on the jet, but not affect other systems until it has been tested and stabilized. In Boeing’s estimation they will be at this point in 3–5 years.

A. RECOMMENDATIONS FOR FURTHER RESEARCH

This research applies specifically to one program. Some recommended areas for further research would address the following research questions:

- How can other specific programs do OT&E in an agile software development environment?
- Once the appropriate infrastructure is in place, does the OT&E process recommended differ in reality? If so, why did the recommendation not work?
- If our recommendations were used for OT&E in an agile software development environment, where can they be improved upon?
- What policies, laws, and/or procedures can be done away with or changed to make OT&E in an agile software development environment more effective?

B. WHY THIS RESEARCH IS IMPORTANT

This research is important in order to get new capabilities into the hands of the warfighter quickly, while still maintaining the trust that the weapons systems they received are effective, suitable, reliable, maintainable, and available as proven by sufficient OT&E. Agile software development is the future, policy makers need to be aware of this now in order to not lose any time deciding how to conduct OT&E in this type of environment. Any delays in these decisions are delays in getting the warfighter what they need to achieve success in the next major conflict with near-peer and peer threats.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Blickstein, I., Yurchak, J. M., Martin, B., Sollinger, J. M., & Tremblay, D. (2016). Navy Planning, Budgeting, Planning, and Execution. RAND https://www.rand.org/content/dam/rand/pubs/tools/TL200/TL224/RAND_TL224.pdf

Brey, D. (2023, April 7). Supersizing SAFe: Enabling Boeing's Digital Transformation. [video]. YouTube. https://www.youtube.com/watch?v=f8jYF_a5YIs&t=1375s.

Director of National Intelligence. (2021, March). The future of the battlefield. https://www.dni.gov/index.php/gt2040-home/gt2040-deeper-looks/future-of-the-battlefield

Government Accountability Office . (2014, March). F-35 Joint Strike Fighter: Problems completing software (GAO-14-322). Government Accountability Office

Government Accountability Office. (1995, June). Impact of China's Military Modernization in the Pacific Region. (GAO 95–84). Government Accountability Office

Government Accountability Office . (2019, March). Including Users Early and Often in Software Development Could Benefit Programs. (GAO-19-136. Government Accountability Office

Government Accountability Office. (2022, February). Challenges Facing DoD in Strategic Competition with China. (GAO-22-105448). Government Accountability Office

NAWCWD (n.d.) | NAWCWD—Naval Air Systems Command. Retrieved January 19, 2023, from https://www.navair.navy.mil/nawcwd/About-NAWCWD

The Office of the Director of Operational Test and Evaluation. (n.d.). The DOT&E mission. https://www.dote.osd.mil/About/Mission/

Office of the Under Secretary of Defense for Acquisition and Sustainment. (2020). Agile software acquisition guidebook. https://www.dau.edu/cop/it/DAU%20Sponsored%20Documents/AgilePilotsGuidebook%20V1.0%2027Feb20.pdf.

Operational Test and Evaluation Force. (2022). Best practices 32—OPTEVFOR process for software acquisition practices. [Memorandum]. 3960 Ser 00TD/50

Operational Test and Evaluation Force. (2020). Operational test and evaluation manual(COMOPTEVFORINST 3980.2J).

OPTEVFOR. (n.d.). Welcome to OPTEVFOR. https://www.cotf.navy.mil/

PMA-265 (n.d.) | Navair—Naval Air Systems Command. https://www.navair.navy.mil/organization/PMA-265

Rendon, R. G., Snider, K. F., & Mortlock, R. F. (2019). Management of defense acquisition projects. American Institute of Aeronautics and Astronautics, Inc.

Under Secretary of Defense for Acquisition and Sustainment. (2020a). Operation of the Adaptive Acquisition Framework [Memorandum]. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500002p.pdf

Under Secretary of Defense for Acquisition and Sustainment. (2020b). Operation of the software acquisition pathway [Memorandum]. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500087p.PDF

Acquisition Research Program
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net