NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

QUANTIFYING THE RISK MANAGEMENT FRAMEWORK

by

Mark I. Heier and Angel J. Morales

June 2020

Thesis Advisor: Dan C. Boger
Co-Advisor: Scot A. Miller

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2020	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE
QUANTIFYING THE RISK MANAGEMENT FRAMEWORK

5. FUNDING NUMBERS

6. AUTHOR(S) Mark I. Heier and Angel J. Morales

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

For the past thirty-five years the DOD/DON have worked diligently to address the exponentially increasing challenges that cyber security presents. While the current Risk Management Framework (RMF) approach improves upon its predecessors, it is once again in need of an overhaul. Derived from National Institute of Standards and Technology (NIST) and DOD directives, the DON's RMF process blindly inherited the ambiguity necessary for larger governing organizations, failing to tailor the RMF to specific Navy organizational needs and practices. The DON RMF is highly qualitative and lacks standardized definitions, measurements, metrics, and a risk assessment methodology. The qualitative approach of the current RMF is further complicated by the bias, heuristics, groupthink, inconsistency, overconfidence, and overestimation ensuing from subjective inputs manifested throughout the DON RMF. The DON RMF must have a more quantitative RMF consisting of standardized definitions, measurements, metrics, and better training to ensure risk is being measured and mitigated appropriately. These improvements would continuously provide feedback for process improvement, leading to increased cybersecurity and resiliency of naval networks.

14. SUBJECT TERMS Risk Management Framework, RMF, National Institute of Standards and Technology, NIST, qualitative, quantitative, cyber security, information systems security, cyber resiliency, risk	**15. NUMBER OF PAGES** 113
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

QUANTIFYING THE RISK MANAGEMENT FRAMEWORK

Mark I. Heier
Lieutenant, United States Navy
BS, Strayer University, 2012

Angel J. Morales
Lieutenant, United States Navy
BS, U.S. Naval Academy, 2013

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN NETWORK OPERATIONS AND TECHNOLOGY

from the

NAVAL POSTGRADUATE SCHOOL
June 2020

Approved by: Dan C. Boger
Advisor

Scot A. Miller
Co-Advisor

Thomas J. Housel
Chair, Department of Information Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

For the past thirty-five years the DOD/DON have worked diligently to address the exponentially increasing challenges that cyber security presents. While the current Risk Management Framework (RMF) approach improves upon its predecessors, it is once again in need of an overhaul. Derived from National Institute of Standards and Technology (NIST) and DOD directives, the DON's RMF process blindly inherited the ambiguity necessary for larger governing organizations, failing to tailor the RMF to specific Navy organizational needs and practices. The DON RMF is highly qualitative and lacks standardized definitions, measurements, metrics, and a risk assessment methodology. The qualitative approach of the current RMF is further complicated by the bias, heuristics, groupthink, inconsistency, overconfidence, and overestimation ensuing from subjective inputs manifested throughout the DON RMF. The DON RMF must have a more quantitative RMF consisting of standardized definitions, measurements, metrics, and better training to ensure risk is being measured and mitigated appropriately. These improvements would continuously provide feedback for process improvement, leading to increased cybersecurity and resiliency of naval networks.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

A&A	Assessment & Accreditation
AI	Artificial Intelligence
AO	Authorizing Official
AOR	Area of Responsibility
ATO	Authority to Operate
C	Compliant
C2C24	Compile to Combat in Twenty-Four (Hours)
C&A	Certification and Accreditation
CBT	Computer-based Training
CCRI	Command Cyber Readiness Inspection
CCORI	Command Cyber Operational Readiness Inspection
CIO	Chief Information Officer
CISQ	Consortium for IT Software Quality
CJCS	Chairman of the Joint Chiefs of Staff
COVID-19	Coronavirus Disease 2019
CNO	Chief of Naval Operations
CNSS	Committee on National Security Systems
CNSSI	Committee on National Security Systems Instruction
COA	Course of Action
CSG	Carrier Strike Group
CSI	Cyber Security Inspection
CSICP	Cyber Security Inspection and Certification Program
CTN	Cryptologic Technician Networks
CTO	Computer Tasking Orders
DC3I	DOD Cybersecurity Culture and Compliance Initiative
DESRON	Destroyer Squadron
DIACAP	Department of Defense Information Assurance `Certification and Accreditation Process

DII	Defense Information Infrastructure
DITPR	Department of Defense Information Technology Record
DITSCAP	Department of Defense Information Technology Certification and Accreditation Process
DOD	Department of Defense
DoDI	Department of Defense Instruction
DON	Department of the Navy
eMASS	Enterprise Mission Assurance Support Service
FCC	Fleet Cyber Command
FISMA	Federal Information Security Management Act
GIG	Global Information Grid
GPEA	Government Paperwork and Elimination Act
GPRA	Government Performance and Results Act
IA	Information Assurance
IPCC	Intergovernmental Panel on Climate Control
ISACA	Information System Audit and Control Association
ISCM	Information Security Continuous Monitoring
IS	Information System
ISO	International Organization for Standardization
ISO	Information System Owner
ISSE	Information System Security Engineer
IT	Information Technology or Information Systems Technician
IV&V	Identify Verify and Validate
KS	Knowledge Service
LP-HC	Low Probability/High-Consequence
MBSE	Model Based Systems Engineering
ML	Machine Learning
NASA	National Aeronautics and Space Administration
NATOPS	Naval Air Training and Operating Procedures
NAVWAR	Naval Information Warfare Systems Command
NC	Non-Compliant

NCDOC	Navy Cyber Defense Operations Command
NIST	National Institute of Standards and Technology
NMCI	Navy/Marine Corps Intranet
NOC	Network Operations Center
NQV	Navy Qualified Validator
OPNAV	Office of the Chief of Naval Operations
PHIBRON	Amphibious Squadron
PM	Program Manager
POA&M	Plan of Actions and Milestones
PPSM	Ports, Protocols, Services Management
PSO	Package Submitting Officer
RAISED	RMF Rapid Assess and Incorporate for Software Engineering in a Day
RAR	Risk Assessment Report
RMF	Risk Management Framework
SAP	Security Assessment Plan
SAR	Security Assessment Report
SCA	Security Control Assessor
SDLC	Software Development Life Cycle
SECDEF	Secretary of Defense
SECDEVOPS	Security and Development Operations
SECNAV	Secretary of the Navy
SME	Subject Matter Expert
SP	Security Plan
SP	Special Publication
SPAWAR	Space and Naval Warfare Systems
SSP	System Security Plan
SYSCOM	Systems Commands
TFCA	Task Force Cyber Awakening
TTP	Tactics Techniques and Procedures
UNSECNAV	Under Secretary of the Navy

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

We would like to first thank our families: Sarah Morales, Alexandria Heier, and children, Roman and Elliana Heier, for their unwavering support throughout our naval careers and completion of our master's degrees.

Next, we want to thank our thesis advisors, Dr. Dan Boger and Scot Miller, for their expertise and guidance throughout the thesis writing process. We would also like to thank several experts who graciously provided their time and subject-matter expertise, allowing us to learn a great deal about the DON RMF Process. Bill Denham, Manfred Koethe, and the entire NAVWAR 5.8 team, we will forever be grateful for your support.

Lastly, we would like to thank Alison Scharmota, from the Graduate Writing Center, and Carla Orvis Hunt, our editor, for reviewing our thesis, assisting us with improving our writing skills, and helping to elevate our thesis to a whole new level.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

For the past thirty-five years, the Department of Defense (DOD) and the Department of the Navy (DON) have worked diligently to address the exponentially increasing challenges that cyber security presents. A pivotal part of information security is the assessment of risk. Is a system safe generally? What are its vulnerabilities? The assessment of potential events is vital to the Navy. While the current Risk Management Framework (RMF) approach improves upon its predecessors, it may again require an overhaul. Derived from National Institute of Standard and Technology (NIST) and DOD directives, the DON's RMF process blindly inherited the ambiguity necessary for larger governing organizations, failing to tailor the RMF to specific Navy organizational needs and practices. The DON RMF is highly qualitative, lacking standardized definitions, measurements, metrics, and risk assessment methodologies. The qualitative approach of the current RMF is further complicated by bias, heuristics, group think, inconsistency, overconfidence, and overestimation ensuing from subjective inputs manifested throughout the DON RMF. The DON needs a more quantitative RMF consisting of standardized definitions, measurements, metrics, and better training to ensure risk is being measured and mitigated appropriately, in order to continuously provide feedback for process improvement, leading to increased cybersecurity and resiliency of naval networks.

A. PROBLEM

The current DON RMF is highly qualitative and lacks standardized definitions, measurements, metrics, and a risk assessment methodology. The qualitative nature of DON RMF is further complicated as subjective inputs are used throughout various parts of its process. These inputs are generally in the form of subjective verbal statements, risk matrices, color coding, and other communications that may introduce errors via bias, heuristics, group think, inconsistency and overconfidence or overestimation of risk into the assessment (Bazerman & Moore, 2013; Fischhoff, 1982; Hubbard, 2009; Hubbard & Seiersen, 2016; Kahneman, 2011; Klipstein, 2017). Often, risk matrices have embedded subjective inputs that become embedded into other risk matrices. At a minimum, risk

matrices have several areas of concern including: a lack of standardization and meaning, range compression, the presumption of regular intervals, and the addition of mathematical errors and inconsistencies (Ball & Watt; 2013; Cox, 2008; Hubbard, 2009; Hubbard & Seiersen, 2016; Klipstein, 2017; Levine, 2012; Peace, 2017). Worse, no evidence exists that the use of risk matrices reduces risk or improves risk management decisions. In fact, the current DON Cyber Security Review of 2019 highlighted that the current approach to cyber security and risk management shows evidence to the contrary. Although there have been numerous recent pushes to make RMF and cyber security faster, such as Compile to Combat in 24 Hours (C2C24) and RMF Rapid Assess and Incorporate for Software Engineering in a Day (RAISED), the DON needs to reprioritize its efforts to accurately account for risk.

B. NEED FOR A QUANTITATIVE MODEL TO MEASURE RISK

The current DON RMF process depends on many variables and uncertainties. Additionally, without a standardized risk assessment model, the resulting approach to risk varies within DON organizations. Without standardization, each DON organization ends up making subjective estimates of risk that can vary widely due to individual bias, heuristics, groupthink, and overconfidence. A quantitative risk assessment would not immunize risk approaches from subjectivity; however, quantitative approaches better account for subjectivity while qualitative risk models do not (Hubbard & Seiersen, 2016). The DOD and DON do not have a standardized risk model, terminology, approach, or means of measuring success. The DON needs a quantitative, multi-criteria risk model that accounts for the aforementioned issues and concerns arising from subjective inputs and risk matrices. At a minimum, a more quantitative model would help reduce uncertainty. Especially in cyber security, even a small reduction of uncertainty can yield significant results (Hubbard & Seiersen, 2016). While the scope of this research pertains to the DON specifically, many of the findings are also be applicable to the DOD as a whole.

C. PURPOSE

The purpose of this study is to analyze the current DON RMF process, risk definitions, approaches, and measurements; to identify potential areas of improvement; and

to propose well-documented, effective risk management methods for immediate consideration within the DON RMF. Because the DON's current RMF process does not contain a means to maintain pace with the ever-changing cyber security landscape and to keep the overall risk manageable, this thesis identifies issues within the RMF and provides a strategy for reducing information system security risk, improving decision-making, and growing operational cyber resiliency.

D. METHODOLOGY AND STUDY DESIGN

We use a qualitative-based research method emphasizing literature review, case studies, and subject matter expert interviews. Our research examines the need for more quantitative approaches and measurements of risk management aimed at effectiveness rather than compliance within the RMF process.

1. Study Design and Implementation

We include comprehensive examination of the Risk Management Framework in our research's qualitative methodology, including examining how the DON defines and measures risks. In various forms of research from April 2019 to March 2020, we studied the following areas:

- Classroom study in pertinent subjects such as Secure Management of Systems, Cyber Security Incident Response and Recovery, Information Sciences for Defense, Computer-based Tools for Decision Support, and Technology Enabled Process Improvement (CS3670, CS4684, IS3001, IS3301, IS4220)

This instruction provided a baseline knowledge on strategy and policy governing the DOD as well as technology and tools utilized in support of Cyber Resiliency.

- Interviews and conversations with Subject Matter Experts (SME) in the fields of Cyber Resiliency, Risk Management Framework, Security Controls, Model Based Systems Engineering (MBSE), and Software Engineering

- Attendance at the Consortium for Information & Software Quality (CISQ) Cyber Resilience Summit in Washington, DC, in October 2019.

This annual summit is convened to brief Federal and State IT leaders and policymakers regarding standards and best practices for measuring risk and quality in software. Various government and industry leaders presented and facilitated panel discussions reinforcing the need to apply proven standards and methodologies that reduce risk and assist in meeting the objectives for acquiring, developing, and sustaining secure and reliable systems.

- Numerous site visits to Naval Information Warfare Systems Command (NAVWAR) in San Diego, California to meet with leading SMEs directly supporting RMF and its processes within the Department of the Navy

- Case study analysis of qualitative and quantitative measures and approaches to RMF within the Department of Defense and leading industry members.

These case studies allow for the juxtaposition of qualitative versus quantitative methods, highlight advantages and disadvantages, and reinforce the need for an approach to RMF that includes valuable aspects of both.

- Case study analysis of cognitive psychology and its effects on the thought process in decision-making

These case studies identified the numerous challenges and considerations associated with subjective inputs and interpretations while highlighting its impacts in qualitative and quantitative methods and addressing the methods and considerations to mitigating these challenges.

- Case study analysis of risk matrices and their impacts within multiple industries ranging from insurance, financial, safety, oil and gas, and cyber security.

These case studies allow for examination of the effectiveness of risk matrices within various fields along with challenges and outcomes associated with their usage while

additionally providing recommendations and rules when risk matrices must be used within certain organizations.

2. Scope and Limitations

While the RMF is a DOD-wide program, this thesis specifically focuses on how to improve the RMF within the DON. Our methodology included researching current definitions and measurements of risk within the DOD, DON, and leading members of industry to evaluate the current RMF within the DON, to identify deficiencies, and to recommend improvements. In 2020, the COVID-19 pandemic restricted our ability to travel specifically to conduct an in-depth analysis of Authority to Operate (ATO) packages, a highly recommended step for future researchers.

3. The RMF Analysis

The preceding methodology informed a deep understanding of the DON's definition and measurement of risk. This thesis hypothesizes that the DON's RMF is qualitative, focused on compliance rather than effectiveness, and that more quantitative measures would better ensure cyber resiliency.

We conducted extensive risk management research via case studies throughout various fields to correlate challenges that persist across the fields as well as key mitigation methods and ingredients for success. The case studies allowed the identification of common mistakes and fallacies within risk management that are widely applicable across industries to include the DON, RMF, and cyber security.

We compared our findings to the DON RMF process and consulted SMEs within the DON RMF. Through these consultations, we confirmed issues within the RMF and filled information gaps between the research and our interpretations.

4. Research Questions and Hypothesis

1. How does DON define risk (cyber)?
2. How does industry define risk (cyber)?
3. How does DON measure risk?

4. How does industry measure risk?

5. Does RMF adequately support cyber resiliency?

Hypothesis: The current state of the RMF is not at a sufficient level to deal with the emerging real-world threats that are posed via cyber security. The DON's approach to risk management is a more qualitative approach focused on compliance rather than effectiveness. Examining successful risk management solutions, we will propose potential, effective risk management methods and metrics for immediate consideration within DON RMF.

E. ORGANIZATION OF THESIS

Chapter I introduces the RMF, the problem and purpose of the research, the methodology and study design, and thesis objectives. Chapter II reviews literature that defines risk and examines potential advantages and disadvantages of the various risk measurement methods and metrics. Chapter III provides step-by-step analysis of the DON RMF, and Chapter IV analyzes the DON RMF in light of Chapter II's best practices. Last, Chapter V provides DON RMF improvement recommendations and proposes further research topics.

II. LITERATURE REVIEW

A Risk Management Framework (RMF) ideally assures information systems security through a process that identifies, selects, and verifies requirements for system authorization while continuously monitoring said conditions throughout the system life cycle. While the first chapter discussed potential problems within the DON RMF current process, Chapter II's review of the literature examines underlying components of the RMF and how they interact. The chapter does so in four main sections followed by a conclusion: first, it defines risk; second, it details underlying principles of methods and measurement and various benefits; third, it analyzes potential advantages and disadvantages of current risk management methods; and, fourth, it examines underlying facets of subjective inputs and estimates and ways to mitigate subjectivity.

A. RISK DEFINED

What is *risk*? While a firefighter might assess running into a burning building differently than would a schoolteacher assess walking into a school, the concept of risk shares a common denominator across industries: the potential for loss and how to mitigate it. Add to the general objective of avoiding loss the significant challenge for decisionmakers who need to balance multiple objectives (Klipstein, 2017), and risk assessment grows more complex. The following sections provide a historical overview and definitions.

1. Historical Assessment of Risk

Humans have long sought to manage risk to avoid hazards or reduce loss. We found the earliest example emerging around 3000 BC when Chinese merchants began spreading their cargo around multiple vessels to minimize potential losses (Vazzano Ltd, n.d.). Around the same time, Babylonians sold insurance in the form of bottomry contracts (Greene, 2018). Under such contracts, merchants would receive loans that would convert to grants if the shipment sunk under the provision that the interest on the loan offset the insurance risk (Greene, 2018). Later, the Hindus in 600 BC and the Greeks in 400 BC continued this practice (Greene, 2018). Insurance contracts grew rapidly through the Greek

and Roman eras, more so as trade and financial activities expanded in Europe in the late Middle Ages, and even further once the Americas entered the picture (Roggi & Altman, 2013). The growing need for insurance created a demand for means to calculate risk, and in 1494, mathematician Luca Pacioli began his work based off of gambling chances (Roggi & Altman, 2013). Several additional scholars continued toward a probabilistic calculus that could help assess event probability. Particularly, Daniel and Jacob Bernoulli provided two powerful theoretical improvements: *the law of large numbers*, which argued that averages obtained during large trials of random variables were more likely to be accurate than those obtained during smaller trials and "subjective elements (risk aversion) in the theory of choice under uncertainty" (Roggi & Altman, 2013, p. 7). Research expanded while these core elements remain the basis of risk assessment approaches today.

2. Definitions of Risk: General and Industry

While etymologists debate whether *risk* can be traced back to the Spanish *risco*, "a pointed, sharp rock" or "dangerous for navigation," the Latin *resecare*, "to cut," or the Arab *ritz,* "all that is necessary to live" (Roggi & Altman, 2013), in simple terms, the word *risk* primarily means the potential for a negative consequence to occur.

Leading industry organizations often influence the development of DOD/DON frameworks; however, industry leaders have more than one universal definition of risk. The Information Systems Audit and Control Association (ISACA) defines risk as "the combination of the probability of an event and its impact" (ISACA, n.d.). The International Organization for Standardization (ISO) defines risk as "the combination of the probability or frequency of occurrence of an event and the magnitude of its consequence" (International Organization for Standardization, 2009). The national leader in standards development, the National Institute of Standards and Technology (NIST), defines risk as "a measure of the extent to which an entity is threatened by a potential circumstance or event, and typically a function of: (i) the adverse impacts that would arise if the circumstance or event occurs; and (ii) the likelihood of occurrence" (NIST, 2012, p. 69).

3. Definitions of Risk, DOD, and DON

Different DOD organizations also utilize varying definitions of risk. The Department of Defense Risk, Issue, and Opportunity Management Guide for Defense Acquisition Programs defines risk as: "a potential future event or condition that may have a negative effect on achieving program objectives for cost, schedule, and performance. Risks are defined by (1) the probability (greater than 0, less than 1) of an undesired event or condition and (2) the consequences, impact, or severity of the undesired event, were it to occur." (Department of Defense, [DOD], 2017, p. 78). The Committee on National Security Systems (CNSS) Publication 4009 glossary defines risk as: "a measure of the extent to which an entity is threatened by a potential circumstance or event, and typically a function of: (i) the adverse impacts that would arise if the circumstance or event occurs; and (ii) the likelihood of occurrence" (Committee on National Security Systems, 2015, p. 104).

B. MEASUREMENT: METHODS, PURPOSE IN DON, AND BENEFITS

In the context of cyber security, some SMEs argue that some things simply cannot be measured, yet we know measurement serves a purpose for the Navy and offers benefits. Hubbard and Seiersen (2016) posit a series of misunderstandings about measurement methods, what is being measured, and even the definition of measurement itself. This section examines those misunderstandings, including Hubbard and Seiersen's three reasons for the common misconception that certain things within cybersecurity and information security are immeasurable: concept of measurement, object of measurement, and methods of measurement. The section then examines the role of performance measurements in the Navy and several benefits of using measurements.

1. Measurement Misunderstandings in Cybersecurity: Concept, Object, and Methods

Definition confusion and variance can give a false appearance of immeasurability. The *concept of measurement* has various interpretations because the definition of measurement appears to be widely misunderstood. Our research did not yield a definition or application of measurement in DOD or DON literature pertaining to RMF; for purposes

of this thesis research, we adopted the proposed definition from Hubbard and Seiersen (2016, p. 21) "a quantitatively expressed reduction of uncertainty based on one or more observations." However, since DON RMF is derived from the NIST RMF, it is important to examine the how measurement is defined by NIST. NIST SP-800-55 Revision 1 (p. 9) defines measurement as "the process of data collection, analysis, and reporting." Similarly, the *object of measurement* is not well-defined further complicating the ability to provide accurate measurements. More complications occur when objects are defined with loose and ambiguous language (Hubbard & Seiersen, 2016). Using well-defined and understood measurements facilitates proper communication regarding the purpose of the measure.

The examination illuminates several other measurement components that support decision-making. For example, according to NIST (2008), there are three areas for measurement: implementation, efficiency/effectiveness, and impact. NIST (2008, p. 13) defines measuring implementation as measures "used to demonstrate progress in implementing information security programs, specific security controls, and associated policies and procedures." Although numerous types of measures can be used concurrently, the information security (IS) measures shift as the IS program matures (NIST, 2008). According to NIST (2008), once all controls, policies, and plans are implemented, measurements will shift towards effectiveness/efficiency and impact. The purpose of effectiveness/efficiency measures are to track whether program-level processes and systems-level security controls are correctly implemented, operated, and meet the desired outcome based on risk assessment results. The purpose of measuring impact is to communicate the effect that information security has on the organization's mission. This varies by organization but is utilized to quantify countless mission-related impacts of information security such as effects on cost, public trust, and many others.

Often, many objects believed to be unmeasurable are not only measurable but perhaps have already been measured (Hubbard & Seiersen, 2016), and unfamiliarity with *methods of measurement* can complicate the situation. Often, cybersecurity experts are unfamiliar with many procedures of empirical observation (Hubbard & Seiersen, 2016). Requirements specified for calculating measurements by NIST are measures that "yield quantifiable information for comparison purposes, apply formulas for analysis, and track

changes using the same points of reference" (NIST, 2008, p. 9). Commonly, cybersecurity experts use percentages or averages or sometimes absolute numbers, depending on the activity they are measuring (NIST, 2008). Most agree that data is necessary to support quantifiable performance measurements. NIST (2008) recommends considering the following when searching for data to calculate performance measures:

- Data for calculating measures must be readily available, and the process needs to be measurable;
- Processes must be consistent and repeatable to be considered for measurement;
- Even in repeatable and stable processes, measurable data might be difficult to obtain if they are not well documented; and
- Easily obtainable data should be used to reduce the resource burden on the organization (NIST, 2008, p. viii).

NIST recommends that the information security activities used to provide quantifiable data for conducting measurements are risk assessments, security assessments, penetration testing, and continuous monitoring (NIST, 2008). The effectiveness of training and awareness programs is also quantifiable and is listed as a successful trait employed by Fortune 500 companies (SECNAV, 2019).

Although NIST emphasizes measurement data being easily obtainable, Hubbard and Seiersen propose additional considerations. Hubbard and Seiersen (2016) believe that individuals claiming that cybersecurity lacks ample amounts of data to conduct quality statistical measurements are generally math-phobic managers and that, in reality, there is more data to measure than one might think. Two examples are using data from other organizations or systems, even if not completely identical to ours, and measuring system components or the whole system (Hubbard & Seiersen, 2016). Hubbard and Seiersen (2016) describe various scenarios such as: How else would a doctor know that a drug that their patient never tried might help them? How can an insurance company estimate someone's health if they have never made a claim? How can an engineer predict a system's

behavior that has not been built yet? Statistics and science would be much easier if everything measured could be directly seen, but most measurements that would be considered "hard" involve indirect deductions and inferences (Hubbard & Seiersen, 2016). These concepts also apply to cybersecurity as system details are derived from many different data sources that include embedded, unseen components (Hubbard & Seiersen 2016).

2. Role of Performance Measurements in the DON

Regulatory, financial, and organizational reasons drive the requirement to measure information security performance (NIST, 2008). Researchers can use performance measures to support internal improvement efforts, linking information security programs to agency level strategic planning efforts (NIST, 2008). For the DON, a number of existing rules, regulations, and laws mention performance measurement generally and Information Security specifically, including the Clinger-Cohen Act, the Government Performance and Results Act (GPRA), the Government Paperwork Elimination Act (GPEA) and the Federal Information Security Management Act (FISMA) (NIST, 2008). According to NIST (2008), the following factors must be considered during the development and implementation of an information security measurement program:

- Measures must yield quantifiable information (percentages, averages, and numbers);
- Data that supports the measures needs to be readily obtainable;
- Only repeatable information security processes should be considered for measurement; and
- Measures must be useful for tracking performance and directing resources (NIST, 2008, p. viii).

NIST SP 800-55 Revision 1, *Performance Measurement Guide for Information Security,* "expands upon NIST's previous work in the field of information security measures to provide additional program level guidelines for quantifying information

security performance in support of an organization's strategic goals" (NIST, 2008, p. 1). The security controls identified in NIST SP 800-53, *Security and Privacy Controls for Federal Information Systems and Organizations,* serve as the foundation to develop measurements that support the evaluation of information security programs. Additionally, NIST SP 800-30, *Guide for Conducting Risk Assessments*, and NIST SP 800-53 are relied upon for supporting quantifiable performance results (NIST, 2008). NIST (2008, p. viii) states that "these measures indicate the effectiveness of security controls applied to information systems and in support of information security programs. Such measures can facilitate risk-based decision-making, improve performance, and increase accountability through the collection, analysis, and reporting of relevant performance related data" (NIST, 2008, p. 1). This allows "the implementation, efficiency, and effectiveness" of an information system and its security controls to be tied to an agency achieving mission success (NIST, 2008, p. 1).

The most recent review of Cyber Security within the DON highlighted a need for RMF and cyber security performance measurements (SECNAV, 2019). The 2019 review of the Cyber Security posture within the DON discovered a "lack of means to adequately measure or even estimate the cost or value of items at risk," which inhibits "the ability to provide justification for investments in this area" (SECNAV, 2019, p. 12). This suggests that the aforementioned NIST publications are not as useful as they should be.

3. The Benefits of Using Measurements

Performance measurements provide key organizational benefits including: "increased accountability, improved information security effectiveness, simplified compliance demonstration, and quantified inputs for resource allocation decisions" (NIST, 2008, pp. 10–11). Utilizing measures that determine if specific security controls are implemented incorrectly, are ineffective, or are not implemented at all, in addition to tracking personnel responsible for implementing those security controls increases accountability (NIST, 2008). Information security effectiveness improves through the quantification of progress toward: organizational goals and objectives; information security processes, procedures, and security controls implemented; and comparing real-

world activities and events (i.e., data breaches, server downtime, etc.) (NIST, 2008). Compliance is easier to demonstrate by having performance measures in place and readily available, which also reduces data collection time.

One of the most compelling benefits of performance measurements is the ability to provide quantifiable inputs for resource allocation decisions. With government and the DON operating on a reduced budget due to fiscal constraints and market conditions, it can be difficult to justify investments into information security. The usage of "information security measures" supports "risk-based decision-making by contributing quantifiable inputs to the risk management process" (NIST, 2008, p. 11). This directly allows organizations to measure their successes or failures of "information security investments while providing quantifiable data to support resource allocation for risk-based decisions" and investments, which yields the best value from available resources (NIST, 2008, p. 11).

Measurements provide many organizational benefits and ultimately aim to improve knowledge through observation-based uncertainty reductions about a quantity that is relevant to a decision (Hubbard, 2009). To observe any measurement, a baseline must be established, and benchmarks must be set for success to be measured (NIST, 2008). This is particularly difficult for an organization like the DON where success is not measured in money, but in the far more nebulous concept of security.

Therefore, the Navy has a crucial need to set performance targets to assist with measurements. Performance targets are a critical component when defining "information security measures as they will establish the benchmark by which success is measured" (NIST, 2008, p. 30). NIST (2008, p. 30) states that "setting performance targets for effectiveness, efficiency, and impact measures is much more complex as management will need to apply qualitative and subjective reasoning to determine the appropriate levels of security effectiveness and efficiency for benchmark performance."

C. RISK MANAGEMENT METHODS

While several different risk management methodologies exist, they all propose to identify risk and reduce it to an acceptable level. This research focuses on the means of identifying those risks, aiming to analyze qualitative and quantitative risk management

methods within the DON RMF. Considering that risks are not always readily apparent, identifying those risks and their sources as early as possible is crucial. The following three sections examine three underpinnings of risk management: risk assessment; qualitative, quantitative, and semi-quantitative approaches; and the special problems of risk matrices.

1. Risk Assessment

Risk assessment is a fundamental component of an organizational risk management process as described in NIST SP 800-39. Risk assessments "identify, estimate, and prioritize risk to organizational operations (i.e., mission, functions, image, reputation), organizational assets, individuals, other organizations, and the nation, resulting from the operation and use of information systems" (NIST, 2012, p. 1). According to NIST SP 800-30 (2012, p. 1),

> The purpose of risk assessments is to inform decisionmakers and support risk mitigations by identifying: (i) relevant threats to organizations or threats directed through organizations against other organizations; (ii) vulnerabilities both internal and external to organizations; (iii) impact (i.e., harm) to organizations that may occur given the potential for threats exploiting vulnerabilities; and (iv) likelihood that harm will occur.

Determination of risk is the end result of a risk assessment ("i.e., typically a function of the degree of harm and likelihood of harm occurring") (NIST, 2012, p. 1).

Risk assessments "are conducted at all three tiers in the risk management hierarchy: Tier 1 (organization level), Tier 2 (mission/business process level), and Tier 3 (information system level)" (NIST, 2012, p. ix). For example, at Tiers 1 and 2, "organizations use risk assessments to evaluate systemic information security-related risks associated with organizational governance and management activities, mission/business processes, enterprise architecture, or the funding of information security programs" (NIST, 2012, p. 1). At Tier 3, "organizations use risk assessments to more effectively support the implementation of the Risk Management Framework" (NIST, 2012, p. 1).

2. Qualitative, Quantitative, and Semi-quantitative Approaches

There are several avenues to approach risk assessment, based on three standard approaches: qualitative, quantitative, and semi-quantitative. NIST SP 800-30 (2012) and

Rot (2008) discuss qualitative, quantitative, and semi-quantitative approaches to risk management, each having specific advantages and disadvantages. The selected approach may vary based on the culture of an organization and their attitude towards the notion of uncertainty and risk communication (NIST, 2012). NIST SP 800-39, *Managing Information Security Risk: Organization, Mission, and Information System View*, states that a risk assessment is a key component of "a holistic, organization-wide risk management process" (NIST, 2011, p. 6). A risk management process includes: (i) framing risk; (ii) assessing risk; (iii) responding to risk; and (iv) monitoring risk (NIST, 2012, pp. 33–45).

Qualitative risk assessments generally utilize non-numeric principles, rules, or scoring methods to categorize risk levels by terms such as high, moderate, or low (NIST, 2012). Qualitative risk assessments easily communicate results to decisionmakers (NIST, 2012). Additionally, they allow risks to be prioritized in order, determining areas with greater risk in less time and for less money (Rot, 2008). Disadvantages of qualitative risk assessments include "disallowance of determination of probabilities and results using numerical measures, cost-benefit analyses being more difficult during the selection of protections, and the results achieved having general characters or approximations since rankings are using an ordinal based measuring system" (Rot, 2008, p. 2).

Additionally, with the small range values common to qualitative risk assessments—low, moderate, or high—it becomes more difficult to prioritize and compare reported risks (NIST, 2012). NIST SP 800-30 states that, "unless each value is clearly defined or characterized, different experts, relying on their individual experiences, could produce significantly divergent assessment results" (NIST, 2012, p. 14). Thus, qualitative risk assessments leave room for ambiguity and subjectivity in the overall determination of risk.

Quantitative risk assessments utilize numerical principles, rules, or methods to estimate risk based on available data (NIST, 2012). The data could be defined in amounts, frequency of threat occurrences, or susceptibility by probable loss value, all of which allow for numerical definitions to estimate the probability of a risk occurrence and whether the potential consequence is acceptable (Rot, 2008). Quantitative risk assessments, therefore, effectively support the cost-benefit analysis of alternate risk responses or courses of action

(COA) (NIST, 2012; Rot, 2008) and gives a more accurate risk picture (Rot, 2008). However, disadvantages include imprecision and confusion with analysis results, lessened accuracy of results influenced by subjective determination, and outweighed benefits affected by cost, expert experience, and analysis tools (NIST, 2012; Rot, 2008). Hubbard & Seiersen (2016) believe that a quantitative approach is best for conducting risk assessments including in the cyber security field.

Combining benefits from both qualitative and quantitative risk assessments, *semi-quantitative risk assessments* utilize a set of principles, rules, or methods to estimate risk with bins, scales, or representative numbers from which their values and definitions are not defined in other contexts (NIST, 2012). Examples could be bins (e.g., 0–15, 16–35, 36–70, 71–85, 86–100) or scales (e.g., 1–10) (NIST, 2012). This combination allows for easier translation into qualitative communications for decisionmakers and value comparisons within the same bin or separate bins (NIST, 2012). Disadvantages include lessening accuracy of results with subjective determinations, poorly defined bin ranges, and experts relying on their individual experiences, which could produce inconsistent assessment results (NIST, 2012).

3. Risk Matrices

Many times, the information required to support the quantification or measurement of risk appears lacking or hard to obtain, resulting in the search for a qualitative approach (Hubbard, 2009; Hubbard & Seiersen, 2016; Wall, 2011), often leading to the use of risk matrices to support a qualitative approach (Hubbard & Seiersen, 2016).

A popular risk-assessment/risk-management methodology tool utilized by a wide array of public, private, and government organizations, risk matrices may even be considered by some as a best practice. Markowski and Mannan (2008, p. 152) define a risk matrix as "a mechanism to characterize and rank process risks that are typically identified through one or more multifunctional reviews (e.g., process hazard analysis, audits, or incident investigation)." Cox (2008, p. 497) defines a risk matrix as "a table that has several categories of 'probability,' 'likelihood,' or 'frequency' for its rows (or columns) and

several categories of 'severity,' 'impact,' or 'consequences' for its columns (or rows, respectively)."

Like in many other fields, prevalent among approaches for cyber security experts is the use of risk matrices based on ordinal scales (Hubbard & Seiersen, 2016). These scales normally represent likelihood and impact with categories labeled as very low, low, medium, high, very high, or sometimes use numbers on a scale of one to five. See Figure 1 for an example of a risk level assessment scale providing a qualitative approach to risk assessment based on the combination of likelihood and impact.

TABLE I-2: ASSESSMENT SCALE – LEVEL OF RISK (COMBINATION OF LIKELIHOOD AND IMPACT)

Likelihood (Threat Event Occurs and Results in Adverse Impact)	**Level of Impact**				
	Very Low	**Low**	**Moderate**	**High**	**Very High**
Very High	Very Low	Low	Moderate	High	Very High
High	Very Low	Low	Moderate	High	Very High
Moderate	Very Low	Low	Moderate	Moderate	High
Low	Very Low	Low	Low	Low	Moderate
Very Low	Very Low	Very Low	Very Low	Low	Low

Figure 1. Assessment Scale – Level of Risk (Combination of Likelihood and Impact). Source: NIST SP 800-30 (2012, p. I-1).

Usually, the matrix's impact axis contains numerical values associated with it in a semi-quantitative fashion. The impact axis is normally based on a qualitative scale, where the levels of impact are judgment-based. However, these scales may have implicit quantitative values that might not be recognized. The likelihood and impact scores are mapped to a cell within the matrix. The cells are generally assigned numbers called *risk scores* and aim to represent a quantitative assessment of the risk with higher scores indicating higher risks.

Other times, rather than assigning specific numbers or scores to cells, a risk matrix shows a color-coding schema assigned to denote the overall score. The color-coding often

appeals as an easy visualization tactic. For example, see how Figure 2's bright red shows a very likely event associated with a very high impact, warning of danger or high risk. The number and color schemes in a risk matrix can be combined (see Figure 3).

Likelihood (Threat Event Occurs and Results in Adverse Impact)	**Impact**				
	Very Low	**Low**	**Moderate**	**High**	**Very High**
Very High	Very Low	Low	Moderate	High	Very High
High	Very Low	Low	Moderate	High	Very High
Moderate	Very Low	Low	Moderate	Moderate	High
Low	Very Low	Low	Low	Low	Moderate
Very Low	Very Low	Very Low	Low	Low	Low

Figure 2. Risk (Combination of Likelihood and Impact). Source: DOD RMF KS, https://rmfks.osd.mil/rmf/RMFImplementation/AssessControls/Pages/ResidualRisk.aspx (accessed 2020).

Matrix		Impact Level				
		1	2	3	4	5
Likelihood	5	9	15	18	23	25
	4	6	12	16	19	24
	3	4	10	13	17	22
	2	2	7	11	14	21
	1	1	3	5	8	20

Figure 3. Example of a Risk Matrix with Numbers and Color Codes. Source: Ibtida and Pamungkas (2018).

Sometimes, the matrix also contains explanations relating to the labels or numbers, specifying the qualitative terms in semi-quantitative values. There may even be additional descriptions in an effort to help better define the qualitive terms. The NIST Special

Publication 800-30 risk level assessment scale shows a qualitative and semi-quantitative approach to risk assessment (see Figure 4).

TABLE I-3: ASSESSMENT SCALE – LEVEL OF RISK

Qualitative Values	Semi-Quantitative Values		Description
Very High	96-100	10	**Very high risk** means that a threat event could be expected to have **multiple severe or catastrophic** adverse effects on organizational operations, organizational assets, individuals, other organizations, or the Nation.
High	80-95	8	**High risk** means that a threat event could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, individuals, other organizations, or the Nation.
Moderate	21-79	5	**Moderate risk** means that a threat event could be expected to have a **serious** adverse effect on organizational operations, organizational assets, individuals, other organizations, or the Nation.
Low	5-20	2	**Low risk** means that a threat event could be expected to have a **limited** adverse effect on organizational operations, organizational assets, individuals, other organizations, or the Nation.
Very Low	0-4	0	**Very low risk** means that a threat event could be expected to have a **negligible** adverse effect on organizational operations, organizational assets, individuals, other organizations, or the Nation.

Figure 4. Assessment Scale – Level of Risk. Source: NIST SP 800-30 (2012, p. I-2).

Unsurprisingly, risk matrices have been highly adopted as approaches to risk management, receiving praise for simplicity and effectiveness; a simple table to assess such complex factors is often welcome. According to Cox (2008), risk matrices have certain advantages:

- Simple appearing inputs and outputs
- May utilize flashy colored grids
- Easy documentation for rationale of risk priority and ranking settings
- Allow systematic review of individual and groups of risks
- Facilitate stakeholder participation to categorize category definitions and action levels
- Allow for organizational training on risk culture at all levels

- Assist leadership in defining categories and expressing risk appetite
- Creation/usage requires no expertise in quantitative risk assessment or data analysis (Cox, 2008, p. 498).

Despite risk matrices' popularity, no evidence actually exists that the matrices reduce or improve risk management decisions. Rather, evidence does show that risk matrices lead to issues: lack of standardization and meaning, range compression, the presumption of regular intervals, and the addition of mathematical errors and inconsistencies (Ball & Watt; 2013; Cox, 2008, Hubbard, 2009; Hubbard & Seiersen, 2016; Klipstein, 2017; Levine, 2012; Peace, 2017). Some issues are easily corrected while remedying others proves more problematic (Thomas, Bratvold, & Bickel, 2013).

Further demonstrating the concern for risk matrices' lack of standardization is an experiment conducted in 2009, which demonstrated that having clearly defined values does not guarantee a universal interpretation. In the experiment, over 200 participants read sentences from the 2007 report on the Intergovernmental Panel on Climate Control (IPCC) and were asked to assign numerical values to probability terms (Budescu, Broomell, & Por, 2009). Definitions of the numerical values were available for terms such as Very Likely, Likely, Unlikely, and Very Unlikely (Budescu et al., 2009). Results of the experiment proved that the participants still subjectively assigned their own values to these terms even with access to the definitions (Budescu et al., 2009). For example, participants interpreted *Likely* to mean between 45% and 84% even though *Likely* was defined as greater than 66% (Hubbard, 2009; Budescu et al., 2009). While one seeming benefit of qualitative risk assessments is improved communication and understanding (Romona, 2011), this experiment suggests the contrary, as this method can cause confusion and errors in communication (Budescu et al., 2009).

Risk matrices demonstrate two additional issues: range compression and the presumption of regular intervals. Range compression results when the same qualitative rating is assigned to vastly different quantitative risks (Cox, 2008), generally because of a large risk range per category (Levine, 2012). For example, consider numbers in a risk assessment with a 1 to 5 or 1 to 10 scale and how large of an impact even minor movement

would have (Klipstein, 2017). Usage of small amounts in the scale for most of the ratings further amplify this problem. For example, Hubbard and Seiersen (2016) examined cyber security risks from five separate organizations and found that roughly 76% of the scores were either a 3 or 4. This clustering of scores can have an even larger impact on risk-based decisions. Most individuals might assume that the numbers on a scale are at least relatively close to the magnitudes associated with those items and that they have regular intervals, if the scale range decreases then the magnitude of impact would increase (Klipstein, 2017). For example, assuming that the scale of 1-2-3-4-5 contains regular intervals would mean that a score of 4 is twice as good or bad as a 2 although this is not necessarily true (Hubbard, 2009; Savage, 2002).

This area of concern with risk matrices and the addition of mathematical errors and inconsistency is highlighted by the extensive research conducted by Cox, Babayev & Huber, 2005; Cox, 2008; Hubbard, 2009; Hubbard & Seiersen, 2016; Thomas, Bratvold & Bickel, 2013. Cox et al. (2005) discusses *uninformative ratings* and *reversed rankings*. *Uninformative rantings* occur when "assigning the most severe qualitative risk label to small quantitative risks while giving the same ranking to risks differing by many orders of magnitude" (Cox et al., 2005, p. 1). *Reversed rankings* occur when "higher qualitative risks are assigned to situations that have lower quantitative risks" Cox et al., 2005, p. 1). Thomas et al. (2013) discuss ranking reversal in their research, which examined 30 risk management case studies in the oil industry. They found that, in 5 out of 30 case studies, the risk scores were reversed. Thus, risk matrices rankings are "arbitrary; whether something is ranked first or last, for example, depends on whether one creates an increasing or a decreasing scale" (Thomas et al., 2013, p. 18). Cox (2008, p. 1) also believes that risk matrices may result in suboptimal resource allocation because "effective allocation of resources to risk-reducing countermeasures are not based on the same categories." The categorization of inputs (i.e., likelihood or impact), and resulting outputs (i.e., risk ratings), requires subjective judgments and interpretations, further complicated by uncertainty (Cox, 2008; Hubbard, 2009; Hubbard & Seiersen, 2016).

D. SUBJECTIVE INPUTS AND ESTIMATES

Many risk-based decisions made by Subject Matter Experts (SME) require subjective inputs and estimates. These subjective judgments, coupled with uncertainty, induce inconsistencies often amplified by the different people involved in the decision-making process and do not account for additional individual heuristics or biases. The following sections examine potential impacts of subjectivity, calibration, and overconfidence, overestimation, and inconsistency.

1. Subjectivity

Uncertainty within risk management clouds consideration and induces risk in the mind of decisionmakers; therefore, in many organizations, including the DON, leaders do not make decisions without SME counsel or advice (Klipstein, 2017). These inputs are generally in the form of subjective verbal statements, risk matrices, color coding, and other communications that may introduce errors via bias, heuristics, group think, inconsistency, overconfidence, or overestimation of risk into the assessment (Bazerman & Moore, 2013; Fischhoff, 1982; Hubbard, 2009; Hubbard & Seiersen, 2016; Kahneman, 2011; Klipstein, 2017).

2. Calibration Applied to SMEs

Professional technicians, engineers, and scientists are trained to only use tools and instruments that are properly calibrated. SMEs must apply a similar rigor when being utilized as tools within risk assessments (Hubbard & Seiersen, 2016). In many risk decisions, SMEs assess probabilities directly or indirectly. They can estimate probability quantitatively, “there is a 10% chance this system will fail”; qualitatively, “it is unlikely this system will fail”; or semi-quantitatively, “on a scale of 1 to 5, the likelihood of this system failing is a 1.” Considering that human experts serve as tools during risk management processes, it would seem reasonable to want to know their performance record in assessing the probability, impact, or even identification of potential risks (Hubbard, 2009; Hubbard & Seiersen, 2016). Human judgment is present in nearly every type of risk management, even the most advanced quantitative analysis of risks (Hubbard & Seiersen, 2016). Research has shown that nearly all people, including SMEs and managers, do very

poorly at assessing the probabilities of future events (Hubbard, 2009; Hubbard & Seiersen, 2016; Kahneman & Tversky, 1972; Keren, 1987).

Even with these issues, research suggests that SMEs and managers should support the risk management process and insist that the system properly accounts and adjusts for the likely shortfalls (Hubbard & Seiersen, 2016); one way to do so is through calibration training. Although SMEs and managers might be experts in their fields, they are almost never experts in estimating probabilities such as impact and likelihood; in other words, unless one is a probabilities expert, one is definitely unlikely to be able to apply probabilities correctly. Training, which Hubbard (2009) refers to as calibration, can help correct this issue. Calibration can be described as the correlation between a person's judgment of execution and their actual execution (Keren, 1991). Research suggests that, when compared, calibrated individuals making probability assessments significantly outperform non-calibrated individuals (Hubbard & Seiersen, 2016; Lichtenstein, Fischhoff, Phillips, 1981). "Assessing uncertainty is a general skill that can be taught with a measurable improvement" (Hubbard, 2014, p. 95). Chapter IV examines if this calibration exists in the DON RMF process, and, if not, where its implementation might assist in improved risk management and increased cyber resiliency.

3. Overconfidence, Overestimation, and Inconsistency

Research also suggests that overconfidence might have the largest impact on the vulnerability of human judgment to biases, particularly because of its commonality and its potential to inadvertently add errors and additional biases (Angner, 2006; Bazerman & Moore, 2013; Fischhoff, 1982; Herrmann, 2013; Kahneman, 2011). According to Prims and Moore (2017), the three types of overconfidence, overestimation, overplacement, and overprecision, are defined as follows:

- Overestimation, the belief that you are better than you truly are at something
- Overplacement, an exaggerated belief that you are better than others at something

- Overprecision, the excessive belief that you know the truth about something (Prims & Moore, 2017, p. 29)

Like most organizations, the DON RMF process is not immune to overconfidence (Gardner, 2010; Silver, 2012). Also, incentives and the level of effort put into identifying potential surprises can increase overconfidence (Lichtenstein & Fischoff, 1977). An additional challenge of measuring overconfidence and overestimation is that time is often a limiting factor. It could take years of collecting estimates from experts to compare them to real world outcomes.

A person who overestimates risk is one who assigns a higher risk value than is necessary (Klipstein, 2017). For example, most people tend to overestimate the probability of rare events happening (Wouter Botzen, Kunreuther, & Michel-Kerjan, 2015) as evident in the key findings of a 2013 research study conducted to determine how well IT experts assess risk (Herrmann, 2013). Unlike other fields such as medicine or natural disasters, where nonexperience often contributes to underestimations and previous experience leads to more realistic probability estimates, previous experiences with IT risks lead to an overestimation of probability (Herrmann, 2013). Decisionmakers often overweigh low probabilities of success while under-weighing higher risks (Klipstein, 2017), which can also be referred to as low-probability/high-consequence (LP-HC) events, or *tail events* (Wouter Botzen et al., 2015). For an everyday example, consider that some people buy lottery tickets but not insurance even though they live in a likely disaster area (Heilbronner et al., 2010; Kahneman & Tversky, 1984; Klipstein, 2017; Woutzer Botzen et al., 2015).

Ramler and Felderer (2013, p. 95) conducted a study involving SME risk estimates of a web-based application. Six experts, heavily involved in the development phase, were asked to provide estimates on the probability of a component or application part being defective. Researchers recorded the risk estimations—given under the subjective ratings terms of high, medium, or low—at the end of the development phase then compared to the data results following the testing phase. SMEs agreed on risk estimates 16.7% of the time when the components were mostly estimated as low probability (see Figure 5). Conversely, SMEs significantly disagreed on risk estimates 16.7% of the time, which, in this context,

means that one expert classified a component as a high-risk probability while another classified the same component as low.

Figure 5. Expert Estimates of Risk Probability for Application Parts and Components. Source: Ramler and Felderer (2013, p. 95).

Overall, the expert estimations had a 52.8% accuracy rate. Of the study's 36 components, SMEs underestimated seven of fifteen high-risk probabilities (see Figure 6). Many components estimated as medium risks were, contrastingly, overestimated, and the remaining components estimated as low risks were slightly overestimated.

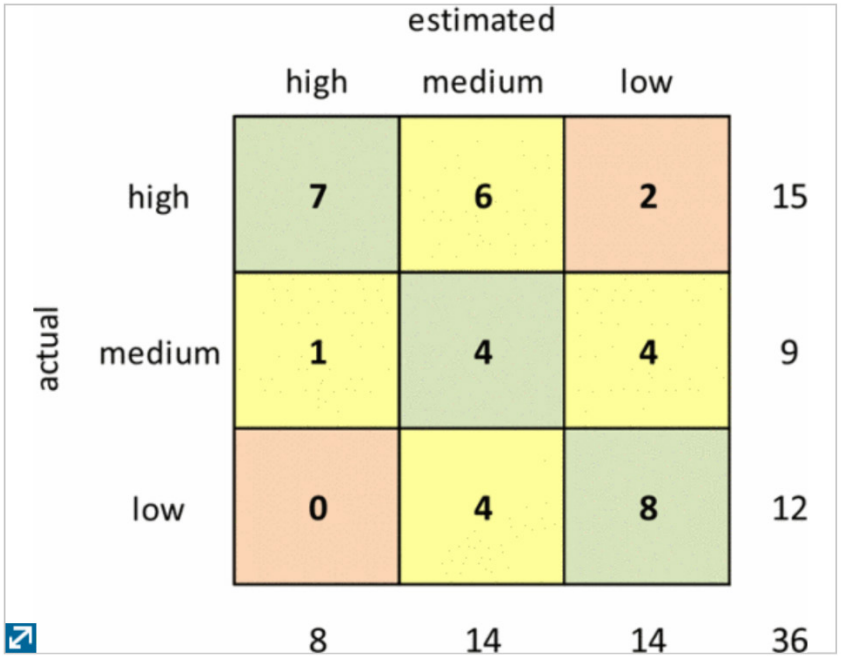

Figure 6. Estimated Risk Probability vs. Actual Risk Exposure. Source: Ramler and Felderer (2013, p. 96).

A different way to measure subjectivity, which is not limited by time in the same manner, is the measure of consistency amongst SMEs. According to Goldberg (1968), consistency can be measured by *stability* or *consensus*. Stability refers to when an expert agrees with their own previous judgments in an identical situation. Consensus refers to when one expert is in agreement with other experts. Researchers have observed experts as inconsistent in both stability and consensus (Hubbard & Seiersen, 2016; Monti & Carenini, 2000). Ross, Kreinovich, & Wu (2000, p. 450), during research supported by NASA to create an automated knowledge-based system from SME elicitations to aid in making similar decisions when SMEs might not be available, recognized similar issues stemming from the expert's internal reasoning and its formalized representation. When SMEs are faced with an inconsistency, they often make changes to their original estimates that can result in completely different adjustments that could be considered random (Ross et al., 2000).

Research shows that overconfidence increases when more information becomes available, when questions become more difficult, or when regular systematic feedback is lacking or influenced by the cognitive style of the expert (Dawes, 1994; Lichtenstein &

Fischhoff, 1977; Lichtenstein, Fischhoff, & Phillips, 1981; Tetlock, 2005). Also, overconfidence could be reduced provided that experts receive prompt, frequent, and unambiguous feedback (Lichtenstein et al., 1980). Apparently, requesting experts to consider reasons that they may be wrong (Koriat, Lichtenstein, & Fischhoff, 1980) or incentivizing for accuracy (Hoelzl & Rustichini, 2005) both can reduce the overconfidence tendency. Contrastingly, informing experts about overconfidence and directing them to be cautious appears to only make a slight difference (Fischhoff, 1982).

A few key findings from Mcbride, Fidler, & Burgman (2012) stated that, while experts possess valuable knowledge, they may require training to communicate their knowledge more accurately. A 2015 case study involving 2,860 intelligence analyst forecasters, 494,552 forecasts, and 344 individual forecasting questions—over a period of three years—showed that calibration training increased the confidence and accuracy of probability assessments for real world geo-political events (Moore et al., 2017). In short, the performance of a well-calibrated analyst can be assessed by the accuracy of their assessments over time (Moore et al., 2017), and prediction is a learned skill and something that can be improved with practice (Frick, 2015, para. 11).

4. Heuristics

Heuristics guide human judgment and decision-making (Bhatia, 2015, p. 232); mental shortcuts that human beings use to process complex events, unusual situations, and everyday information they encounter (Colwell, 2005). Humans use heuristics to make "educated guesses" when information is lacking (Davis, Kulick, & Egner, 2005; Dowd, Petrocelli, & Wood, 2014; Kahneman, 2003). Basing decisions, often unknowingly, on mental shortcuts can lead to serious errors in logic and reasoning in certain decision-making scenarios (Colwell, 2005). According to Tversky and Kahneman (1974), the three categories of heuristics are: representativeness, availability, and anchoring and adjustment. Finucane, Alhakami, Slovic, & Johnson (2000) added a fourth category: affect. The effects of heuristics can be further compounded by assuming that the human mind is analogous to a machine regarding how risk information may be handled (Joffe, 2003).

The *representativeness heuristic* can explain numerous findings in the judgment and decision-making literature, often notably in the context of risk and uncertainty (Krawczyk & Rachubik, 2019). According to Davis et al. (2005), representativeness happens when someone believes that an object belongs to a class based on how much it resembles that class. A piece of information is considered to have high representativeness if it is judged to have many similar characteristics of a category's salient features (Olson, 1976). This can also apply when comparing new information to other pieces of information. Representativeness can impact how people judge a cause and effect based off the similarity of other observed events (Davis et al., 2005). Tversky and Kahneman (1974) showed that, when estimating probabilities, where representativeness is present, prior probabilities will be neglected for consideration. Managers are often required to make decisions in conditions where information is limited, and they will apply heuristics to fill in information or data gaps (Brookins & Ryvkin, 2014; Wickham, 2003). As demonstrated in the Budescu et al. (2009) case study, representativeness could result in people overriding information that is present and replacing it with their own values, leading to incorrect assumptions that relationships exist between pieces of information when they do not. The representativeness heuristic exemplifies "how practical rules for decision-making may lead to suboptimal decision outcomes" (Wickham, 2003, p. 163).

The *availability heuristic* suggests that the ease with which an event can be brought to mind or imagined is used as a substitute for probability, meaning that, the easier it is to retrieve a similar event from memory, the more likely we are to assume it has a higher probability (Davis et al., 2005; Juslin, 2013; Tversky & Kahneman, 1974). Tversky and Kahneman (1974) gave the example of overestimating the risk of dying from a heart attack in middle age by recalling the occurrences amongst acquaintances: a person not remembering such occurrences may consider the overall probability low-risk and vice versa. Similarly, people are more likely to recall relatively recent events than older events.

Like all heuristics, the probability heuristic is not inherently negative; it can be a useful guide of probability in some real-life circumstances; however, because the ease of retrieval or imagination is affected by factors other than probability or frequency, this heuristic may produce a number of biases or cognitive illusions in probability judgment

(Juslin, 2013; Tversky & Kahneman, 1974). These biases can arise as either a consequence of biases in the external flow of information or because of the properties of encoding and retrieval relevant to human memory (Juslin, 2013). Highly emotional or interesting events tend to receive higher priority and are better encoded in memory (Juslin, 2013). Additionally, imaginability plays an important part in evaluating real world situations, as a disaster may be easier to visualize yet its likelihood may be overestimated, or conversely, a hard-to-imagine event may cause a risk to be underestimated (Tversky & Kahneman, 1974). While the availability heuristic is intended to simplify the situation, it can also limit the consideration of outcomes, a situation that can be further complicated or even detrimental when the decisionmaker has incomplete data (Klipstein, 2017).

The anchoring and adjustment heuristic refers to when a person makes an estimate by starting from an initial value that is then adjusted up or down to arrive at a final answer (Tversky & Kahneman, 1974), normally most evident in negotiations. In the phenomenon known as anchoring (Davis et al., 2005; Epley, 2013; Tversky & Kahneman, 1974), the initial value serves as a starting point and "may be suggested by a formulation of the problem or a result of a partial computation" (Tversky & Kahneman, 1974, p. 1128). Regardless of how the mind establishes the anchor, the adjustments—shiftings from the initial value to reach the final value (Davis et al., 2005)—are generally insufficient as "different starting points will produce different estimates that are biased towards the initial values" (Tversky & Kahneman, 1974, p. 1128). In most situations, the anchor results in a biased final value as people tend to stay close to the anchored value whether it was relevant to the judgment or not (Epley, 2013). The anchoring and adjustment heuristics' effects are compounded when decisions are made quickly (Yudkowsky, 2008) and even more so when estimates are presented as confidence intervals (Tversky & Kahneman, 1974). Numerous studies have asserted that when participants state that they are 98% confident that a number is within a certain range, they are only correct about 60% of the time (Lichtenstein, Fischoff, & Phillips, 1977; Tversky & Kahneman 1974). Negotiation and dispute resolution scholars have identified the profound impact of the anchoring and adjustment heuristic, including Orr & Guthrie's 2006 extensive meta-analysis that showed potential dangers that anchoring had on final outcomes.

In 2000, Finucane et al. proposed that affective feelings may affect judgment and decision-making. Distress, disgust, sadness, anger, fear, surprise, or happiness can all affect decisions (Clore & Huntsinger, 2007; Klipstein, 2017), a phenomenon known as the affect heuristic. Specifically, depictions of events and objects end up tagged with various affects in people's minds to various degrees (Slovic, Finucane, Peters, & MacGregor, 2007) creating an *affect pool* that holds the positive and negative tags associated with conscious and unconscious representations during the process of making judgments. People then make recollections from their *affect pool* (Slovic et al., 2007). Finucane et al. (2000, p. 3) states that affect plays an important role in judgments similar to how "imaginability, memorability, and similarity serve as cues for probability judgments in the availability and representativeness heuristics." According to Zajonc (1980), affective reactions are usually an automatic, initial reaction that guides information processing and judgment. Zajonc (1980) believes that all perceptions contain some affect, giving the example of people, when choosing a car or house, justifying the choice to buy the one they find attractive due to various other reasons. Finucane et al. (2000, p. 3) believes that using an affective impression can be "much easier and more efficient than weighing pros and cons or recalling examples from memory, particularly when making complex decisions where mental resources are limited."

5. Bias

Biases are tendencies to view information from a perspective that prevents the person from being objective and impartial and can lead to poor judgments and decisions (Pfleeger & Caputo, 2012). *Cognitive bias* is referred to as a systematic pattern of deviation from a standard of rationality in judgment (Hastelton, Nettle & Andrews, 2005) and can serve as an error in judgment and decision-making caused by misconceptions (Phillips-Wren, Power, & Mora, 2019). Cognitive biases are present when making judgments and decisions and apply to both everyday individuals as well as experts (Keil, Depledge & Rai, 2007). Previous research has shown that humans tend to make suboptimal decisions from a rational viewpoint (Phillips-Wren et al., 2019). These suboptimal decisions stem from the close connection that cognitive biases have with decision-making because people understand and foster predictable thinking patterns (Dvorsky, 2013; Tversky & Kahneman,

1974). Biases can stem from previous experiences, emotional responses to stimuli, views, beliefs, individual limitations of mental processing power, or perceptions of the environment (Davis et al. 2005; Tversky & Kahneman 1974). Biases are utilized in everyday choices. For instance, we may be biased when deciding what school to send our children to, what kind of suit or dress to buy, who we want to win the Super Bowl, or what color to paint our houses (Klipstein, 2017). Most significantly, bias is present in the heuristics used to foster understanding and perceptions for experts, analysts, and decisionmakers (Tversky & Kahneman, 1974). Many biases can lead to poor estimations of probabilities, frequencies, values, and estimates including, IT-related risk assessments (Hermann, 2013). While literature has not defined a finite list of every bias that every expert, analyst, or decisionmaker may face, the following section discusses some common biases: confirmation bias, attention bias, belief bias, and the clustering illusion (Davis et al., 2005; Gilovich, Vallone, & Tversky, 1985; Klipstein, 2017).

Confirmation bias "is the tendency to acquire, interpret, favor, or recall information in a way that is consistent" with or strengthens "a person's preexisting beliefs" or desired end state (Allahverdyan & Galstyan, 2014, p.1). Confirmation bias can impact how people determine which information they avoid or approach, process and interpret, and recall from memory (Mothes, 2017), and confirmation biases pervade many fields such as psychology, economics, media, politics, and scientific practices (Allahverdyan & Galstyan, 2014). For example, as Mothes (2017) describes, a person who supports raising the minimum wage will most likely read an article that emphasizes the positive aspects of raising the minimum wage, reaffirming the person's existing beliefs. Confirmation bias works the other way, also: that same person would be less inclined to read about the troubles raising the minimum wage could cause (Mothes, 2017). In both ways, confirmation bias can cause detrimental consequences as a person may downplay or ignore relevant information may leading to flawed analysis and suboptimal decision-making.

Attention bias is the tendency for a person to focus their attention towards certain stimuli while ignoring or disregarding other information as irrelevant (Cherry, 2020, para. 2). Attention bias not only impacts the information perceived in the environment, but also the decisions made based upon those perceptions (Cherry, 2020, para. 1). For example,

previous research has shown that smokers will fixate on information or cues pertaining to cigarette smoking (Begh et al., 2013). Attention bias may lead decisionmakers to ignore or downgrade relevant information, resulting in inaccurate or poor choices.

Belief bias is the propensity to assess the strength of argument based on the believability of its outcome, rather than the logical support offered (Sternberg & Leighton, 2004). A person is more likely to accept an argument when the conclusion coincides with their values, beliefs and prior knowledge while dismissing alternative arguments (Evans, Newstead & Byrne, 1993, p. 118). Markovits and Nantel (1989) indicated a significant belief bias in subjects during the production and evaluation of conclusions where belief was predominantly used to resolve uncertainties. According to Davis et al. (2005), there is a broad consensus to the presence of belief bias, that once people make an interpretative story, all subsequent observations will be processed within an interpretive filter where they notice supportive data and discard other data. Therefore, belief bias can lead decisionmakers to fail to process pertinent information because they have already made up their mind, which may result in less than favorable and potentially detrimental choices.

The *clustering illusion bias* is a natural human tendency to see patterns when no patterns actually exist (Gilovich, 1991; Gilovich, Vallone, & Tversky, 1985). Sometimes referred to as the gambler's fallacy (Gilovich, Vallone, & Tversky, 1985; Oskarsson, Van Boven, Mcclelland & Hastie, 2009; Tversky & Kahneman, 1974), clustering illusion bias can be observed when people interpret patterns or trends in available data (Albert, Bedek, Huszar & Nussbaumer, 2017). Gilovich, Vallone, & Tversky (1985) researched basketball games with several NBA and college teams where people believed players to have "hot hands" or to be "streak shooters." The belief in detecting streaks in random sequences is caused by the general misinterpretation of chance according to which brief random sequences are believed to be highly representative of their developing process (Gilovich, Vallone, & Tversky, 1985). They concluded that the biased belief in a "hot hand" was not only erroneous, but could also be very costly (Gilovich, Vallone, & Tversky, 1985). Consider the example of flipping a fair coin 10,000 times (Klipstein, 2017). Streaks of heads or tails will occur; at some points, those sequences might appear as a pattern. The fairness of the coin may be questioned despite Bernoulli's Law of Large Numbers

accounting for the streaks (Oskarsson et al. 2009). The clustering illusion has been observed in judgments made by everyday individuals and experts in numerous fields (Fischhoff, Slavic, & Lichtenstein, 1982; Kahneman, Slavic, & Tversky, 1982; Nisbett & Ross, 1980; Tversky & Kahneman, 1983). This bias could appear in any situation where an analyst or decisionmaker may view patterns as obvious on the surface, but no data exists that support their perception (Klipstein, 2017).

6. Group Dynamics

In the DOD, organizational staff members routinely conduct analysis for the decisionmaker, an approach also prevalent within the DON. Simultaneously, assertive and sometimes aggressive personalities commonly fill key positions to maximize the likelihood of mission success (Klipstein, 2017). Within groups, decisions are rarely a product of an individual (Adams, 2015, para. 1). According to Greer, Caruso and Jehn (2011), in a study analyzing 66 teams in the financial and telecommunications industries, teams with high power personalities had greater levels of internal process conflict and reduced team effectiveness. Subordinates or peers may disagree with one another, or with decisions, to the point that they begin questioning the knowledge, abilities, and understanding of the task at hand (Klipstein, 2017). This cognitive doubt within the disagreeing individuals may then turn into an unspoken agreement once a group decision is made (Asch, 1955, 1956; Gilovich, 1991). Pressures such as time constraints, budgets, or majority consensus can also result in groupthink (Katopol, 2015), which normally results in suboptimal decisions as in the Bay of Pigs (Janis, 1972). According to Bang and Frith (2017), groupthink can result from lack of independence within the group (i.e., the group members are too similar in knowledge, education, and experience, or when members are initially dissimilar, they adapt to one another through social interaction). Groupthink and subjectivity arising from group dynamics may result in the presentation of a flawed analysis to the decisionmaker.

E. CHAPTER CONCLUSION

Risk itself means the potential for loss, and, naturally, different people estimate potential risks of various endeavors differently, leading to an increased need for a shared language and a deep understanding of the underpinnings of risk assessment including

examining whether an idea that something is immeasurable actually is true or resulting from a faulty premise. Of several approaches to risk assessment, individuals often, particularly in the cyber security field, choose a qualitative approach, generally preferred because of its perceived ease of communication and understanding. However, research has overwhelmingly shown numerous issues with qualitative risk approaches, including subjective inputs, which can give rise to overconfidence, heuristics, and bias including groupthink. Issues with qualitative risk approaches are further compounded by the common use of risk matrices, which inject issues such as range compression, a presumption of regular intervals, and mathematical errors and inconsistencies into the risk process. Across the RMF process, such issues make a qualitative approach untenable. While quantitative models are not necessarily immune to similar issues, they do offer a path to address the concerns of subjective inputs and mitigate the challenges that risk matrices pose. For example, individuals may undergo calibration training to improve their subjective estimates and increase consistency. Chapter III examines the current DON RMF process in light of these findings.

THIS PAGE INTENTIONALLY LEFT BLANK

III. RISK MANAGEMENT FRAMEWORK (RMF)

For almost fifty years, the Department of Defense (DOD) has worked on system accreditation and authorization to ensure systems cybersecurity. The DOD accreditation and authorization process is based on "testing and certification for information systems and an assessment of the security risks posed when connecting those systems to the DOD's Global Information Grid (GIG)" (Valladares, 2013, p. 7). In 1972, DOD released the first iteration: DOD Directive 5200.28, Security Requirements for Automated Information Systems; in 1983, DOD released the second: Directive 5200.28-STD, which was then updated in 1988. Through these three directives, DOD formed the foundation for its information system testing and accreditation. Like most foundations, the directives were a solid start yet contained gray areas open to interpretation, particularly regarding processes, which, in turn, led to each military service promulgating its own similar, but separate, guidance for its own accreditation process with similar issues.

Initially, DOD system accreditation processes focused on discrete, individual systems when conducting system tests. In 1997, DOD announced the DOD Information Technology Security Certification and Accreditation Process (DITSCAP), with an objective "to establish a DOD standard infrastructure-centric approach that protected and secured the entities comprising the Defense Information Infrastructure (DII)" (Department of Defense, [DOD] 1997, p. 4). The DITSCAP presented rules and regulations that standardized the Certification and Accreditation (C&A) and contributed "to more secure system operations and a more secure DII" (DOD, 1997, p. 4). The process considered "the system mission, environment, and architecture while assessing the impact of the operation of that system on the DII" (DOD, 1997, p. 4). DITSCAP intended to standardize the Certification and Accreditation (C&A) process across the DOD enterprise architecture and end systems while minimizing risks to an acceptable level. Although DITSCAP ensured the DOD was trending in the right direction, shortcomings remained. Despite systems being recognized as a piece of the larger enterprise architecture, they were still tested and accredited individually (Valladares, 2013). Additionally, DITSCAP was paperwork-intensive and lacked a standardized list of controls.

In 2006, to fix the remaining issues, especially the need for viewing the larger architecture as a whole (Williams and Steward, 2007). DOD supplanted DITSCAP with provisional guidance: the DOD Information Assurance Certification and Accreditation Process (DIACAP), which was finalized in 2007 (Valladares, 2013). Its purpose was to "address the paradigm shift in IA security from an individual information system-level approach, to a DOD-wide enterprise approach of securing information systems in a net-centric environment and for supporting the implementation of IA security during a system's life cycle" (Williams and Steward, 2007). DIACAP consisted of three essential elements to support the conversion of DOD Information Systems to Global Information Grid (GIG) standards and a net-centric environment that enabled assured information sharing in accordance with federal standards and guidelines (Williams and Steward, 2007).

The DIACAP policy accomplished several goals. It defined standards for security and outlined process requirements for the identification, implementation, and validation of IA controls (Williams & Steward, 2007). The DIACAP Knowledge Service (KS) provided a central location for execution and implementation guidance. The Enterprise Mission Assurance Support Service (eMASS) introduced an automated web support tool designed to assist the DIACAP process by reducing paperwork, limiting resource usage, lowering time consumption, incorporating an enterprise perspective, and improving operational sustainment (Department of Defense [DOD], 2007). Overall, DIACAP provided a noticeable improvement to DITSCAP; however, while the DOD fully adopted DIACAP, the remainder of Federal Government and the Intelligence Community used different C&A processes and security controls, and those differences posed significant challenges to interconnectivity and interoperability.

In 2014, to address that issue, DIACAP morphed into the RMF, derived from the National Institute of Standards and Technology (NIST) RMF. The term "morphed" rather than "replaced" is used to demonstrate the promise that the RMF evolved from the same C&A process and controls of DIACAP and would remain applicable throughout the whole Federal Government while allowing reciprocity for interagency applications. Even though a close correlation exists between the intent and processes outlined in the DIACAP and RMF, the RMF meant major changes. The RMF included a stronger integration effort with

the Systems Development Life Cycle (SDLC); a focus on reciprocity between Federal Government systems; continuous Federal Information Security Management Act (FISMA) reporting; and continuous monitoring (Williams & Steward, 2007). Additionally, the NIST SP-800-53 standardized control set replaced the DODI 8500.2, and C&A changed to Assessment and Accreditation (A&A) (Valladares, 2013).

As helpful as DOD's advances in the last fifty years have been for ensuring information security, DOD and DON may need an instruction that enables better cyber resiliency and allows more rapid evolution to match the rapidly changing IT landscape. Building on Chapter II's review of risk and risk assessment methods, Chapter III specifically examines the DON's RMF process through the six current steps.

A. CURRENT DON RMF PROCESS

The DON RMF is a six-step process to provide authorization of IT systems and services (US Fleet Cyber Command (FCC)/Space and Naval Warfare (SPAWAR) Command, 2015). This risk-based cybersecurity approach primarily focuses on managing risks for Navy IT by implementing tailored security controls and incorporating security into all facets of a systems life cycle rather than trying to add it on after the fact (US FCC/ SPAWAR, 2015). The following sections further describe each step (see Figure 7).

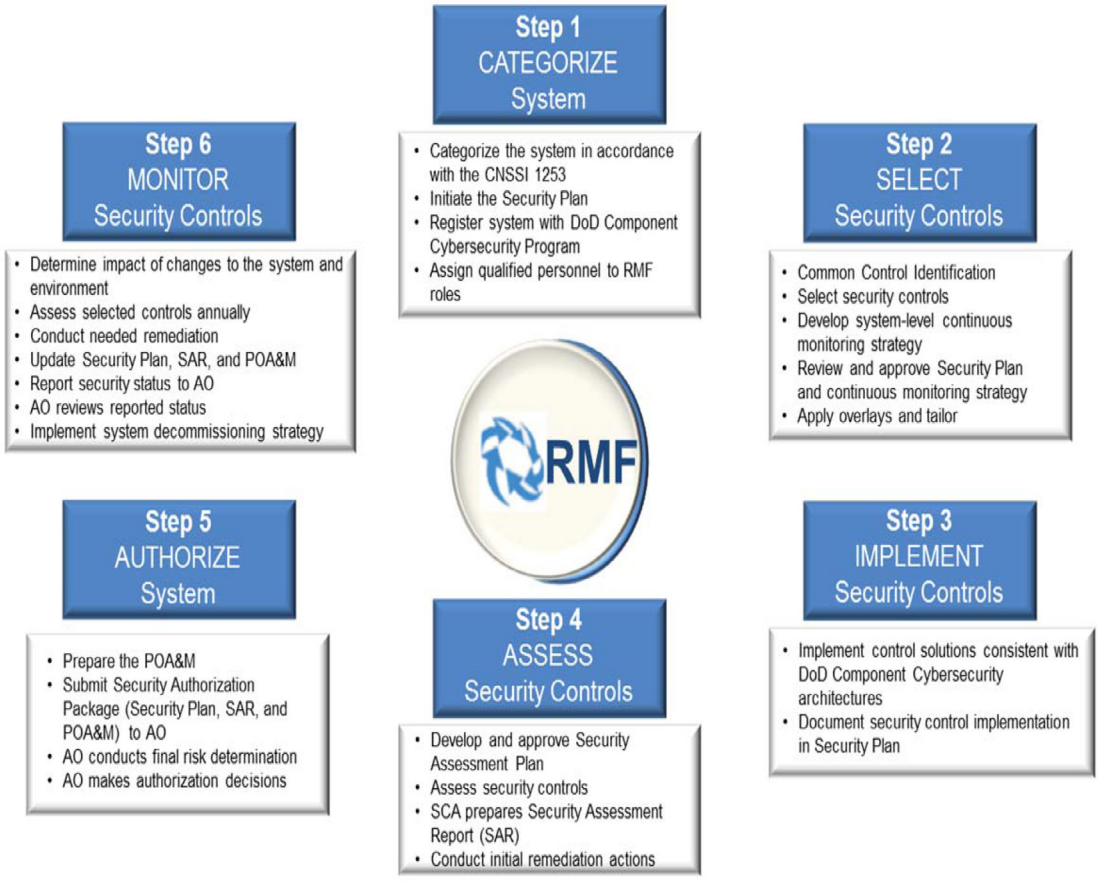

Figure 7. Navy RMF Process Overview. Source: Barrett (2017).

1. RMF Step One: Categorize System

The first RMF step (see Figure 8), is solely administrative and requires the concurrence of all stakeholders to produce a consolidated effort in system categorization. Then, to accurately categorize the system, all stakeholders must define the system boundaries and identify and document all types of information processed, stored, or transmitted by the Information System (IS).

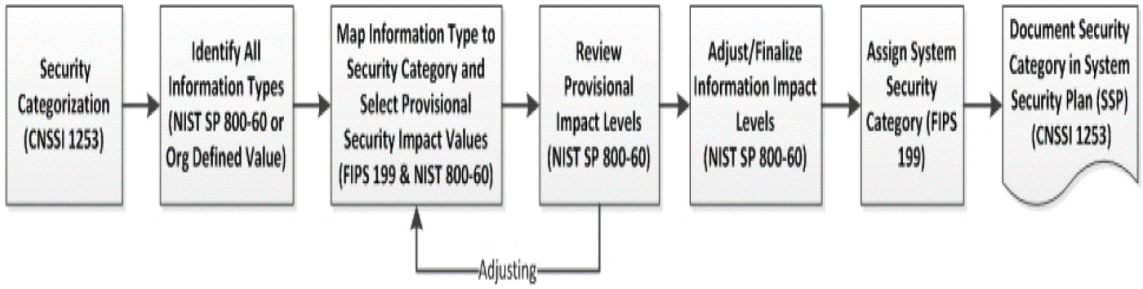

Figure 8. RMF Step One – System Categorization. Source: Barrett (2017).

The primary goal in RMF Step One is to identify potential security impacts if a security breach results from a loss across each of three system security objectives: confidentiality, integrity, and availability. The potential security impacts are classified as either low, moderate, or high (see Figure 9).

POTENTIAL IMPACT[14]			
Security Objective	**LOW**	**MODERATE**	**HIGH**
Confidentiality Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. [44 U.S.C., SEC. 3542]	The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Integrity Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. [44 U.S.C., SEC. 3542]	The unauthorized modification or destruction of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Availability Ensuring timely and reliable access to and use of information. [44 U.S.C., SEC. 3542]	The disruption of access to or use of information or an information system could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

Figure 9. RMF System Categorization Impact Values. Source: Miller, Kiriakou, and Hilton (2015).

At the completion of Step One, a completed Navy System Categorization form, an initial System Security Plan (SP), the beginning of an enterprise Mission Assurance Support Service (eMASS) record, and, if applicable, a DOD Information Technology Portfolio Repository DON (DITPR-DON) record are produced. Based on Step One's system categorization, RMF Step Two identifies the system's security controls.

2. RMF Step Two: Select Security Controls

Once the Information Systems Security Engineer (ISSE) identifies the system's security controls, common controls, and any potentially inheritable controls in Step Two, they would then use applicable overlays to further tailor the system (see Figure 10).

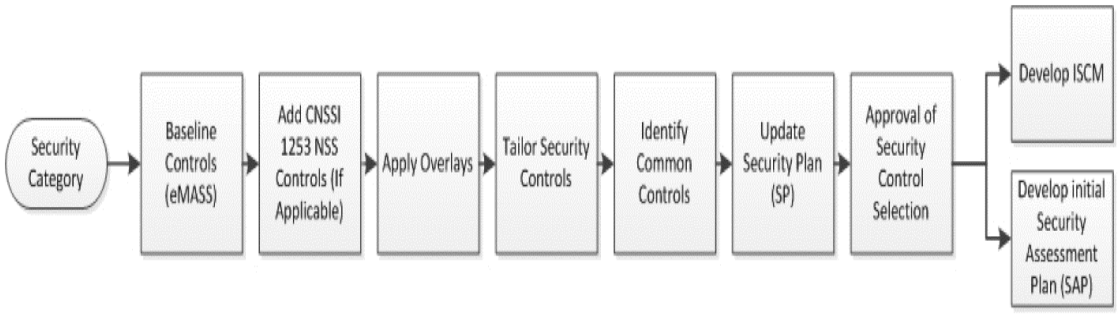

Figure 10. RMF Step Two – Select Security Controls. Source: Barrett (2017).

Then, the Information System Security Manager (ISSM) would update the Security Plan (SP) to identify any non-applicable control tailoring. The Program Manager (PM), working with the ISSM, would develop an initial System Life cycle Continuous Monitoring (SLCM) strategy. The Navy Qualified Validator (NQV), with the support of the ISSE and ISSM, would develop the Security Assessment Plan (SAP) to assess the resulting applicable security controls.

The completion of RMF Step Two requires scheduling a mandatory checkpoint unless waived at the joint discretion of the Authorizing Officer (AO) Cyber Security Analyst (CSA) and the Security Control Assessor (SCA) Liaison. At the Step Two checkpoint, the AO CSA and SCA Liaison would review the SP, Security Assessment Plan (SAP), and SLCM Strategy and either endorse the system for AO approval or return the

artifacts to the originator for remediation. Step Two produces an approved System Categorization form, an updated SP that includes an overview of inheritance relationships, the initial SLCM Strategy, an initial Ports, Protocols, and Services Management (PPSM) Registration form, and the Security Authorization Package.

3. RMF Step Three: Implement Security Controls

In RMF Step Three (see Figure 11), the ISSE, identifies, updates as needed, and implements the security controls and overlays. The ISSE uses the SAP to identify, verify, and validate (IV&V) the security control implementations. Pending the results of the preliminary control testing, if the system's security posture is deemed insufficient, the ISSE then remediates vulnerabilities and retests. When testing verifies that the system's security posture is sufficient, the ISSE documents the security control implementation plan and generates a Risk Assessment Report (RAR) to identify any findings that will not be corrected prior to the formal assessment in Step Four.

Figure 11. RMF Step Three – Implement Security Controls. Source: Barrett (2017).

RMF Step Three produces an updated SP based on the final implementation of tailored security control sets and the results of the security control implementation test performed to verify that the required security controls are in place. The ISSE initiates an RAR and creates, as required, a Plan of Actions and Milestones (POA&M) to correct weaknesses or deficiencies.

4. RMF Step Four: Assess Security Controls

In RMF Step Four (see Figure 12), a Navy Qualified Validator (NQV) performs the official security assessment on behalf of the Security Control Assessor (SCA). The NQV

verifies that each security control has been implemented properly and determines each to be either Compliant (C) or Non-Compliant (NC). Then, based on the assessment results, the NQV would develop the Security Assessment Report (SAR).

Figure 12. RMF Step Four – Assess Security Controls. Source: Barrett (2017).

RMF Step Four's output is derived from the findings of the security control assessment as well as any completed remediation actions. The SP is updated to reflect the true state of the security controls following the initial assessment. The NQV adds any false positives, misleading results, and testing exceptions discovered in the assessment results to the SAP for completion and SAR preparation. The ISSE remediates or mitigates any NC controls or vulnerabilities on the SAR and updates the RAR and POA&M.

Finally, the NQV updates the SAR based on the updated RAR and POA&M in eMASS. Then the ISSM submits the Security Authorization Package and any applicable artifacts to the Package Submitting Officer (PSO) for review and also uploads the same into eMASS for adjudication by the SCA and Authorizing Official (AO).

5. RMF Step Five: Authorize System

In RMF Step Five (see Figure 13), once the package is processed, the PSO schedules the mandatory RMF Step Five Checkpoint, and the AO CSA submits the package to the AO for an authorization decision.

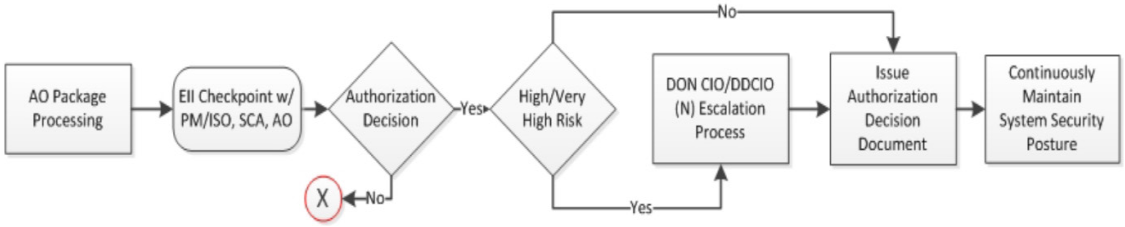

Figure 13. RMF Step Five – Authorize System. Source: Barrett (2017).

The PSO reviews, updates if necessary, and submits the package for review. The Step Five checkpoint must be scheduled; however, it can be waived at the joint discretion of the AO CSA and the SCA Liaison. The NQV finalizes the SAR for review and signature by the SCA with an overall recommendation of system cybersecurity risk. Then the SCA signs and completes the SAR with a statement of overall system cybersecurity risk. The SAR, along with the rest of the package are reviewed by the AO. If the AO determines the overall risk level exceeds the acceptable threshold, then the package will either be returned for remediation or enter the Navy escalation process. If the package represents a compliant system and the overall risk level is determined as acceptable, the AO will issue an authorization decision document. If controls exist with a risk level of "Very High" or "High" that cannot be immediately mitigated or corrected but mission criticality justifies the overall system risk, then, only with the permission of the DON Chief Information Officer (CIO), will an AO issue an Authority to Operate (ATO) with conditions.

6. RMF Step Six: Continuous Monitoring

In RMF Step Six (see Figure 14), the ISSM implements a continuous monitoring strategy as agreed upon in the final authorization decision.

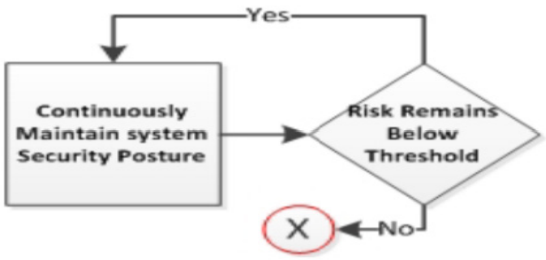

Figure 14. RMF Step Six – Continuous Monitoring. Source: Barrett (2017).

The ISSM tracks the SAP and its associated security controls in eMASS for compliance for the life cycle of the system. Re-authorization occurs as required for system updates or periodicity. This phase is comprised of two subprocesses: first, an annual security review, monthly scanning and patching, and triennial reauthorization; and, second, a conditionally-based subprocess comprised of incident response, Communications Tasking Orders (CTO), Cyber Security Inspection (CSI) and Command Cyber Readiness Inspection (CCRI) findings and decommissioning. If, at any point, events occur that require corrective actions or the system's level of risk exceeds its tolerance threshold, the PM must take necessary corrective actions. If the PM/ISO cannot mitigate the risk level to an acceptable tolerance level, they will request a new assessment for authorization, if applicable, and the system will reenter the RMF process.

B. CHAPTER CONCLUSION

RMF has undergone numerous iterations over thirty-five years to reach its current state within the DON. While it remains a lengthy and complicated process, it also remains crucial to ensure the security of our information systems within today's interconnected world. At each of the six steps, individuals complete specific protocols, leading from stakeholder identification to system reviews. The next chapter applies Chapter II's findings to the RMF process outlined in Chapter III to examine the potential need for a quantitative RMF consisting of standardized definitions, measurements, and metrics, as well as better training to ensure risk is being measured and mitigated appropriately.

IV. ANALYSIS

This chapter applies Chapter II's understandings regarding risk to the current DON RMF process outlined in Chapter III to determine any specific shortcomings and areas for improvement. It does so in five sections followed by a chapter conclusion: impact of varying definitions of risk; impact of varying measurement methods; impact of qualitative versus quantitative risk assessment approaches; specific impacts within the current DON RMF process, step by step; and the impact of five Navy-wide challenges. For assessing best practices, Chapter IV adopts Chapter II's findings: risk assessment works best when: a) terms are well-defined, b) processes are objectively measured, and c) risk assessors are well-qualified. The chapter also identifies procedures where automation and other techniques may increase efficiency of RMF application.

A. STANDARD RISK DEFINITION

The absence of a clear, unambiguous definition of risk—pertaining to cybersecurity in general and the RMF in particular—is troubling. Defining risk is as challenging today as it was when the concept first emerged thousands of years ago, and, without a clear definition, the term is inherently subjective and, therefore, open to interpretation. In short, the lack of a specific risk definition is its own risk. For risk managers to effectively communicate, amongst themselves and with other pertinent stakeholders, they must speak a common language using a shared lexicon. Additionally, since risk management shares similar concepts and requires close coordination with various other fields such as decision analysis, engineering, acquisition, and statistics, the use of differing or unclear terminology adds unnecessary confusion and ambiguity. Improving the communication and understanding of risk factors requires clear terminology in order to reach optimal risk-based decisions.

Further, DON classifies cybersecurity risk in three categories: System Cybersecurity Risk; Operational Cybersecurity Risk; and Residual Cybersecurity Risk (Office of the Chief of Naval Operations [CNO], 2018) but has not tailored these definitions or defined them specifically in the DON RMF Process Guide or other governing

RMF documents. Lack of specifics here makes it more difficult to measure risk and leaves room for subjective interpretation.

B. STANDARD MEASUREMENT METHODS

Similar to risk, the concept of measurement also suffers from varying interpretations and is often misconstrued. NIST SP 800-55 Rev 1 does provide guidance on the measurement of Information Systems; however, despite the instruction being released over 12 years ago, it has not been incorporated into policy regarding measurement within the DON RMF. The RMF aims to reduce risk to an acceptable level, yet the DON RMF gives no deliberate measurements to observe whether risks are, in fact, being reduced. Before anything can be measured within RMF, DON must define a measurement methodology with specific metrics.

The SECNAV Cybersecurity Readiness Review (2019) highlighted the lack of mechanisms to measure or estimate the value of items at risk and their impact within the DON (p. 12). NIST SP 800-55 recommends measuring implementation, effectiveness/ efficiency, and impact. Implementation is intended to focus on ensuring that security controls, policies, and procedures are implemented as required. Once implementation is considered to be 100%, the focus is shifted towards effectiveness/efficiency and impact measures. The DON RMF discusses measures of impact in Steps One through Six and measures of effectiveness and efficiency in Steps Three through Six. However, the DON RMF does not present explicit directions on how to conduct them.

Measurements that do occur within the DON RMF occur in the lab and not in the operational environment. No defined measurements exist in any DON RMF documentation to determine if risk mitigations remain effective and at an acceptable level throughout the life cycle of the system. The failure to take deliberate measurements in the operational environment directly contrasts with the primary goal of RMF, which is to reduce and maintain risk at an acceptable level. All other system performance measures are tested in an operational environment to determine if they are operationally effective and suitable. In our research, it is not clear that the RMF process achieves that standard for operational cyber security and resiliency.

Systems do not always function in operation as they did in a lab environment. Not only does this prevent DON organizations from being able to justify investments to reduce risks, it also suggests there are no means to observe whether risks are being reduced or decisions are being improved under the current state of DON RMF. The words performance, measurement, or any reference to NIST SP-800-55 are rare, if not non-existent, within DoDI 8510.01 as well as the DON process guide. DoDI 8500.01 mentions that performance will be measured and assessed for effectiveness; however, there is no DON guidance on how to accomplish this or any guidance that clarifies the term "effectiveness" within the RMF.

If DON RMF guidance adopted standardized risk methods and risk measurements and placed quantified values on the components of risk, the ability to identify risks and determine objectively if risks are reduced would increase. Adopting such standards would also reduce the level of uncertainty across various risk concerns where even a small reduction of uncertainty within cyber security has significant impact (Hubbard & Seiersen, 2016). As highlighted in the SECNAV CS Readiness Review, "many policy and funding decisions do not reflect the current risk profile" (SECNAV, 2019, p. 28). Having the appropriate measurements in place would allow for the proper allocation and justification of resources and investments as needed, leading to a more effective RMF.

C. QUALITATIVE VERSUS QUANTITATIVE

Additionally, because organizations use different types of approaches—qualitative, quantitative, and semi-quantitative—to assess risk, further confusion ensues. The guiding methodology for conducting risk assessments, NIST SP-800-30, recommends either a qualitative, quantitative, or semi-quantitative approach for each organization, further stating that each approach may vary based on an organization's culture and attitude toward the concepts of uncertainty and risk communication. Compounding the problem, different approaches leave even more room for interpretation within the DON RMF since the Navy has not addressed this ambiguity.

Examination of the RMF process and supporting documents identifies that the current posture of the DON RMF is more qualitative than quantitative. Its qualitative nature

is quite apparent in the risk matrices. The DON RMF appears to be shifting towards a more semi-quantitative approach, utilizing a combination of qualitative descriptions and numerical values to further define associated risk levels. However, the Navy must take care to prevent the inadvertent addition of errors based off subjective inputs, range compression, presumption of regular intervals, and mathematical errors. Other alternatives may allow for more accurate risk assessments and a significant reduction in uncertainty.

Currently, many different personnel provide qualitative inputs to risk matrices. Within the DON RMF, risk matrices are often imbedded into other matrices as well. For example, risk matrices in eMASS come from additional imbedded matrices that identify and select likelihood and residual risk after the risk factors have been defined (Personal Communication, Denham, March 18, 2020). These inputs come from SMEs in their applicable areas of expertise.

Currently, the NQV provides the most inputs for risk matrices; however, the SCA liaison and the SCA are often also involved in this process. The NQV is an independent third party but acts as an extension of the SCA office, validating implementation of the approved security control baseline (Barrett, 2017). The NQV acting as an extension is most observable in its determination of likelihood and risks that depend upon the selected risk factors. The determination of likelihood and risk differ among SYSCOMs because of the subjectivity and a lack of standardization within the DON.

D. DON RMF PROCESS

As long as the Federal Government uses the NIST RMF as the base of their RMF, the DON must incorporate and adapt changes more rapidly. The current DON Process Guide does not address all changes made to the most recent NIST SP 800-37 Rev 2, published in 2018. For instance, NIST 800–37 Rev 2 (2018) expanded upon the previous RMF process by adding a seventh step: Prepare. The following analysis of each DON RMF step identifies current issues and areas where changes are necessary for process improvement, starting with actually including Step Zero.

1. RMF Step Zero: Prepare

Introduced in the most recent revision of NIST SP 800-37 (Revision 2), NIST recommends a preparatory step to ensure organizations are equipped to conduct the six main steps of the RMF (NIST, 2018, p. 8). The purpose of Step Zero is to "complete essential tasks at the organization, mission and business process, and IS levels to help the organization prepare its management of security and privacy risks while utilizing the RMF" (NIST, 2018, p. 11). The prepare step can assist with quantitative measurements provided that standardized definitions, measurements, and metrics exist for the RMF. At the time of this analysis, May 2020, the DON's RMF Process Guide does not include the prepare step.

2. RMF Step One: Categorize System

The categorization step impacts the remaining steps of the RMF process and has a large impact on budget allocation. This step is time-consuming because it requires concurrence between stakeholders and SMEs. That agreement leads to a system impact assessment, which results in the categorization of risk to the Confidentiality, Integrity, and Availability domains as either High, Moderate, or Low. The categorization requires the use of subjective estimates that may vary based on the interpretation of terms and the organization's understanding of the RMF. The efficiency of this step may also vary based upon the organization's approach and understanding of RMF. A standardized and more quantitative model, that builds on the prepare step, would better facilitate RMF Step One. Correct categorization of the system is crucial as it not only affects the remaining RMF steps, but also, if the categorization of risk is done incorrectly, it generates rework in Step Three.

3. RMF Step Two: Select Security Controls

In Step Two, an ISSE selects a baseline of security controls from the RMF Knowledge Service (KS) based on the overall system categorization from Step One. As the primary reference used when selecting security controls, the KS contains a current list of baseline NIST security controls as well as common DOD controls from which the ISSE selects and combines with proper overlays to begin the tailoring process. Since the tailoring can be completed differently by ISSEs, as long as the ISSE justifies their actions, the

selection of overlays and tailoring differences risks influence by the ISSE's subjectivity. Further, the same ISSE may not always select the same overlays or tailor the system the same when dealing with similar or identical systems. To reduce errors within security control selection, the DON RMF process guide states that the ISSE can consult the NQV early in the process to avoid incorrect security control assignment and reduce possible later rework. However, this patchwork fix adds to the NQV workload and also induces validity and consistency questions concerning the ISSE's choices. To avoid NQV increased workload and to still reduce rework costs, a tool that captures the applicability of security controls, overlays, and aids the ISSE in proper system tailoring may prove quite valuable.

4. RMF Step Three: Implement Security Controls

The implementation phase verifies that the tailored security controls and overlays specified in the SP are in place through a preliminary test to gauge the system's security posture in order to identify and remediate any vulnerabilities prior to Step Four. The ISSE must conduct the preliminary test using tools specified in the SAP; however, the ISSE sometimes uses other tools with prior approval from the SCA. The ISSE records the SAP results in eMASS, identifying and measuring the implementation of security controls. This is currently a quantitative measure within the RMF. The current DON RMF Process Guide also states that any vulnerabilities discovered that cannot be corrected prior to Step Four must be annotated in the RAR and POA&M. However, the NAVWAR Introduction to the Risk Management Framework Course, a computer-based training (CBT), states that the RAR is no longer required or used, and this information is now recorded in eMASS. Organizational changes that deviate from written guidance add confusion and can result in more unnecessary risks. This step should be updated to reflect a standardized RMF process.

5. RMF Step Four: Assess Security Controls

A standardized organizational risk assessment methodology, as recommended by NIST, ought to support Step Four. The security controls assessment, conducted by the NQV, verifies that security controls are in place and compliant as specified in the SAP, as well as identifying residual risks for each Non-Compliant (NC) control and vulnerability. Step Four also uses several risk matrices throughout the risk assessment of various system

components as well as the overall system. Absent a standardized organizational risk assessment methodology, Step Four happens in a highly qualitative manner containing subjective inputs from the NQV to identify the severity level of each NC control.

Crucially, the SCA currently uses the risk assessment to identify the overall system cybersecurity risk, risk mitigations, and residual risk within the system as a basis for their recommendation of approval or denial to the Authorizing Official (AO). The DON RMF Process Guide states that the NQV will coordinate with the SCA Liaison and act as their trusted advisor in all matters of risk. The SCA liaison and the SCA spot-check the NQV's results as they have limited time to personally perform all the risk assessments. While the NQV, SCA, and SCA liaison are Cyber Security SMEs, none of them receive specialized training in risk, increasing the potential occurrence of many of the previously discussed issues with risk matrices and subjective inputs. If the risk assessment does not capture the most realistic risk posture, is done incorrectly, or is affected by errors, then the information presented to the authorizing official in Step Five will be already skewed.

6. RMF Step Five: Authorize System

The DON RMF process guide states that the AO will perform a final risk review "to determine the risk to Naval organization operations, organizational assets, individuals, other organizations, or the nation" (Barrett, 2017, p. 56). However, there are numerous authorizing officials within the DON RMF process. Different AOs are found at each Navy Systems Command (SYSCOM), each with its own staff and internal RMF processes. The AOs often have competing responsibilities and are also not SMEs in risk or risk assessment. The AO examines the system's current security state based on the risk assessment and SAR recommendations from Step Four, in addition to any applicable risk-related guidance from senior DOD officials, for "a final determination of risk to DOD operations, assets, individuals, other organizations, and the nation from the operation and use of the system" (Barrett, 2017, p. 55). Examination of the DON RMF process and its supporting documentation leads us to believe that, rather than the AO completing a risk assessment as required in Step Five, the AOs usually simply concur with the SCAs' recommendations; the research could find no examples to the contrary.

If the AO fails to actually perform a risk assessment but instead simply signs off on the NQV and SCA's assessment, large problems can occur. Notably, the NQV and SCA are focused on the system's performance in the lab rather than in the operational environment; the AO's Step Five assessment is meant to consider how a system would actually do operationally. If a senior decisionmaker fails to consider operational performance, in terms of security and risk implications, and, instead, simply concurs with lab findings, Step Five becomes a compliance check, an exercise in checking the box, rather than an actually effective measure to determine risk prior to system authorization.

7. RMF Step Six: Monitor Security Controls

In addition to the very large problems stemming from Step Five "check the box" issues, Step Six, continuous monitoring, further endangers the Navy's risk assessment plan because it also lacks clear definition and specific guidance. Similar to overall risk assessments, continuous monitoring is often interpreted differently from person to person and organization to organization. Interpretation can be especially problematic here since the DON relies on continuous monitoring for ongoing risk assessments once a system is operational. Here, a risk assessment plan that relies on subjectivity both in its construction and in its implementation essentially means there is no plan.

The future goal, even though it has been an actual standing requirement since 2014, is to use continuous monitoring for granting system authorizations since Step Six is a continuous assessment of the system's security controls over its lifetime. In addition to ongoing system scans and network testing conducted by entities such as the Naval Cyber Defense Operations Command (NCDOC), inspections such as the Cyber Security Inspection and Certification Program (CSICP), Command Cyber Operational Readiness Inspection (CCORI), and penetration testing, conducted throughout the fleet, must be incorporated to better facilitate continuous monitoring. The current RMF does not support seamless integration of these results into one source for use in continuous monitoring. When these inspections identify concern for the RMF process requiring further investigation, these concerns should feed directly to the SCA. Unfortunately, since these reports do not feed into this step, they do not seamlessly make their way to the SCA or

even interact with the DON's eMASS tool (Personal Communication, Denham, March 18, 2020).

Automation would better facilitate Step Six; however, currently only approximately 25–30% of security controls appear capable of being automated (Personal Communication, Denham, March 18, 2020). At a minimum, automation would allow other inspections and data about the system to be ingested and viewed by the SCA and would be fed directly to the AO. Instead, verifying that policies and procedures are updated remains a slow and manual process. Ongoing efforts to investigate how Artificial Intelligence (AI) and Machine Learning (ML) might help better facilitate this step should continue as Step Six needs updates, definitions, clarity, and the much-needed help of automation.

E. NAVY-WIDE CHALLENGES LIMITING RMF

Many root causes of risk assessment issues stem from culture, people, structure, process, and resource challenges present throughout the DON today. The following section discusses those five specific DON challenges that limit the RMF and that must be addressed to improve and maintain true cyber warfighting security and resiliency.

1. Culture

The need for a cyber cultural change within the DON is not new; however, translating this idea into noticeable action remains a challenge. The SECNAV Cybersecurity Readiness Review characterized the DON CS culture by "distrust, a lack of knowledge or accountability, a willingness to accepts unknown risks to mission, a lack of unity of effort, and an inability to fully leverage lessons learned at scale" (SECNAV, 2019, p. 14). Since the implementation of the DON RMF, senior leadership has acknowledged that risks are growing based on increasing technological sophistication and the Navy's dependency on connected capabilities and information to fight and win (SECNAV, 2019). The Navy has implemented many initiatives designed to address increasing risk; however, those initiatives have not effected meaningful change to the DON CS culture, which seems to lack "a real appreciation of the cyber threat" (SECNAV, 2019, p. 12).

For example, in 2014, after an adversary breach of the NMCI network, the DON responded with Task Force Cyber Awakening (TFCA). TFCA called for "cultural and organizational changes to meet the increasing threat" and made cybersecurity the "business of every Commanding Officer, calling for an 'All Hands-on-Deck' effort" (Office of the Deputy Chief of Naval Operations for Information Dominance [OPNAV N2/N6], 2014). A year later, the Secretary of Defense (SECDEF) and the Chairman of the Joint Chiefs of Staff (CJCS) established the DOD Cybersecurity Culture and Compliance Initiative (DC3I), which called for a shift in behavior and cultural norms from senior leadership down to the unit level and each individual's cybersecurity performance and accountability (Mauck & Pashley, 2016, para. 8). Despite this, the DON RMF still lacks standard and measurable objectives from which leaders could drive organizations to transition to a system focused on institutional initiative rather than compliance, to increase CS efforts, and to improve risk management (SECNAV, 2019). An RMF that remains geared towards compliance of regulations rather than effectiveness will only provide minimum security, wholly inadequate for overall organizational protection. The 2019 SECNAV Cybersecurity reviews highlights the problem of a "check the box" mentality rather than one truly assessing risk: "the DON cybersecurity governance structure is characterized by an almost exclusive focus on compliance metrics based on a snapshot in time, which inhibit the ability to anticipate and/or adapt to current and future threat environments" (SECNAV, 2019, p. 29). So far, senior leadership's attention to necessary cultural change has not affected a permanent cultural change toward true risk assessment.

2. People

People are obviously integral to innovation and the making of decisions within an organization; however, their actions and decisions expose them to great risk including within the DON RMF process. People, as the Navy's greatest asset and, sometimes, the greatest liability, require tough training, calibration, evaluations, and proper employment. The world's top organizations balance this contradictory point about the value of people by trusting their people while continuously testing and monitoring them to ensure the maintenance of established standards (SECNAV, 2019).

Examination of the DON RMF process and the supporting documentation suggests that individuals assessing risk are not risk experts but rather experts within their respective disciplines including cyber security, financial, and engineering. To become a certified risk assessor within DON RMF requires the completion of six CBT courses. However, none of these six courses pertains to risk, but rather they focus on RMF terms, so that the courses essentially train assessors on how to use an inherently subjective system without training them on the fundamentals of subjectivity; the use of terms without the understanding of those terms and other underlying factors does not help as much as it should. Additionally, each course examination is multiple choice with a required passing score of 80% and unlimited attempts to pass. This short process focused on RMF terminology contrasts with other fields like insurance, safety, and the financial sector where actuaries conduct risk assessments following many years of formalized training including a degree, professional certification tests, and field experience. Considering that the concept of risk management stems from these industries, this research effort finds the lack of similar qualifications and scrutiny in the Navy is particularly alarming since Navy risk assessment failure can lead directly to warfighting causalities.

The DON RMF also apparently lacks any type of calibration training. While the SMEs consulted during this research represent a limited sample size, the authors believe that this finding is applicable to the entire DON. SMEs are employed as the measuring tools within the DON RMF process and therefore require calibration. Given the preponderance of SME input in the DON RMF process, we find the lack of calibration concerning. This concern is further compounded due to the usage of risk matrices imbedded into other risk matrices when making determinations of likelihood, severity, and risk. Case studies ranging from the 1970s to present demonstrate that calibration training yields a remarkable improvement in consistency and accuracy where implemented.

The DON RMF's first five steps require subjective estimates, which often result in numerous sources of error inadvertently added into the Risk Management process. The lack of consistency in using subjective estimates, shown by the review of the current DON RMF process furthers our concern. The DON RMF has no means to account for the potential issues caused by subjectivity, bias, heuristics, or groupthink when estimating risk.

Accounting for these known issues could be greatly helped by a standardized, more quantitative risk model.

Analysis of the DON RMF produced inconclusive results when attempting to determine whether overconfidence and estimation errors are present among risk assessors because no metrics are used and no data is maintained regarding how assessors make estimates. Rather, the NQV ranking system assesses each validator on their ability to conduct risk assessments based on a subjective four-star ranking system from the SCA (Qualification Standards, Responsibilities, and Standards for Navy Qualified Validators (NQV), March 4, 2016). The lack of metrics to capture how assessors make estimates also makes it impossible to assess SME consistency.

The Navy implementing the Identify-Recruit-Train-Sustain-Retain model explained in the 2019 SECNAV Cybersecurity Readiness Review would assist with fixing many people issues present in CS and the DON RMF. The review calls for fostering a career path progression that builds upon cybersecurity expertise throughout the fleet but appears more geared towards the strategic level.

While this call is valid, it does not address the fundamental problem that plagues the Navy: the need for network operators at the operational and tactical levels to assist in continuous monitoring while fighting our Navy systems. The Navy enlisted ITs employed on ships must maintain, monitor, and fight the network; however, most ITs are not trained to monitor the network and the RMF or to fight the ship through a cyber-attack. The Navy does have an enlisted rating who are trained as cyber warriors to fight the network, Cryptologic Technician Networks (CTN). While addressing the issues called out in the SECNAV Cybersecurity Readiness Review, the authors believe that the IT rating also requires a training overhaul to include RMF, continuous monitoring, and fighting of naval networks. Otherwise the CTNs that possess these abilities must be added to ships as well as other operational and tactical units. Absent this immediate adaption, the good intentions of RMF and cyber resiliency are likely to fail.

From a personnel perspective, the Navy's current RMF process currently relies on SMEs who are ill equipped to evaluate risk because they are unarmed with any measurable

facts and are not educated in risk, consistency, or receive calibration training. Further, the Sailors who bear the brunt of the continuous monitoring role lack the training, education, tools, or access to successfully execute these roles.

3. Structure

The current organizational structure within the DON does not foster effective execution of RMF since there is no single authority responsible and accountable for managing RMF across the DON. Importantly, the SECNAV Cybersecurity Readiness report observed that "authority, responsibility, and decision-making power for information risk management has been confused by its distribution within the DON, resulting in fragmented, or uncertain response to the expanding threat" (SECNAV, 2019, p. 28). Under current DON organization, the CIO role is performed by the Under Secretary of the Navy (UNSECNAV), who delegates the responsibility and execution of the role to a Vice Admiral, the N2N6 (SECNAV, 2019). The N2N6 is not a direct subordinate of UNSECNAV but a subordinate of the Chief of Naval Operations (CNO) and the Vice CNO. This structure violates successful industry leaders' best practices since the actual person executing the CIO role does not report to senior executive leadership, in this case, the SECNAV, further reducing authority to completely control the DON CS and RMF (SECNAV 2019).

The delegation of CIO authorities and accountabilities from the UNSECNAV to the uniformed military service deputies is unnecessarily convoluted. Not only does the delegation of authority fail to align with the May 2019 Executive Order, Enhancing the Effectiveness of Agency Chief Information Officers (SECNAV, 2019), it slaps handcuffs on the authority to enforce effectiveness. Absent a well-established "Go/No Go" criteria to defend against continuous threat vectors into naval networks, the current approach becomes merely another "check the box" to appease compliance. For example, a Fleet Commander can tell a ship, aircraft, DESRON, PHIBRON, or CSG that they are not permitted to get underway or fly if they have not met established criteria because they control those assets. However, under the current organization of CIO responsibilities, the N2N6 is not in charge of operational units and does not have the authority to directly

prohibit operations. The DON will not truly be serious about CS until some office/agency within the DON is able to prevent a unit from getting underway due to poor CS hygiene.

4. Process

The Navy is a process-oriented organization, employing some very recognizable processes, such as the Naval Air Training Operations Procedures and Standardization (NATOPS), the Navy Occupational Safety and Health Program, and now the DON RMF, derived from the industry-leading NIST (SECNAV, 2019). Although the DOD's RMF instruction intends to direct a structured cybersecurity process, the September 2018 Cybersecurity Strategy identified specific shortfalls, also cited in the SECNAV Cybersecurity Readiness Review, that contribute to DON processes remaining ineffective in staying ahead of threats.

To further demonstrate how the lack of a standardized risk assessment process results in specific shortfalls within the RMF Process Guide and supporting documentation, these researchers created Figure 15, demonstrating how the current documentation enters a circular loop that lacks specific guidance when conducting risk assessments.

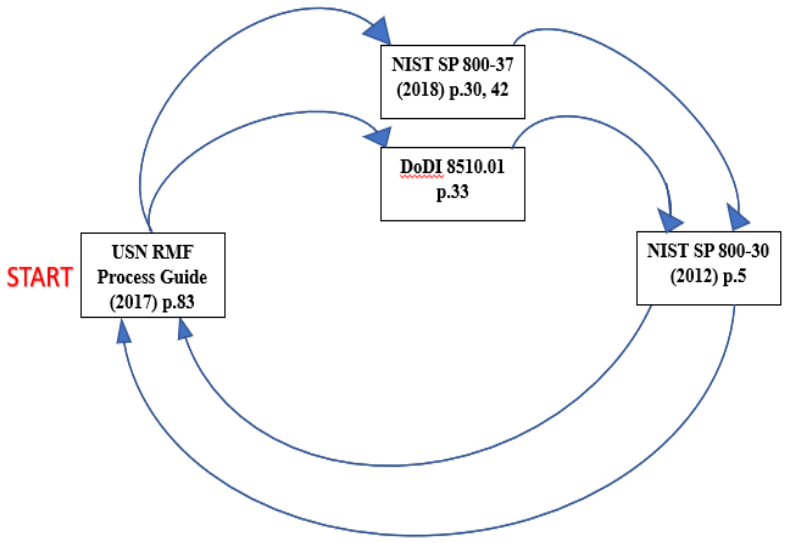

Figure 15. Circular Loops within DON, DOD, and NIST Governing RMF Documentation.

The starting point for any DON risk assessor should be the DON RMF process guide. The DON RMF process guide points to the DoDI 8510.01 and NIST SP 800-37. NIST SP 800-37 (p .42) states that "organizations determine the form of risk assessment conducted (including the scope, rigor, and formality of such assessments) and method of reporting results." Thus, the DON RMF process guide is supposed to be responsible for determining the form of risk assessment conducted. Additionally, NIST SP 800-37 simultaneously points to NIST SP 800-30 for further guidance on conducting risk assessments along with DoDI 8510.01. NIST SP 800-30 (2012, p. 5) states that "risk assessment methodologies are defined by organizations and are a component of the risk management strategy developed during the risk framing step of the risk management process." Both NIST SP 800-37 and NIST SP 800-30 place the organization as the responsible party to determine the type of risk assessment it conducts. However, a standardized or defined risk assessment methodology remains non-existent within the DON or DOD.

Processes for information sharing are also affected by the absence of a standardized process. NIST SP 800-53 states that "realistic assessments of risk requires an understanding of threats to and vulnerabilities within organizations and the likelihood and potential adverse impacts of successful exploitations of such vulnerabilities by those threats" (NIST SP 800-53, p. 5). However, information sharing is inhibited across the DON, other services, and agencies due to the existence of vertically stove-piped organizational structures (SECNAV, 2019, p. 29). When a system is developed and authorized, its set of security controls is based off the intended system's overall purpose and operational use. However, Navy systems operate throughout many different Areas of Responsibility (AORs) around the world, with many different threats and adversarial Tactics, Techniques, and Procedures (TTPs), and, when authorized for use, the set of security controls assigned do not address every threat in every AOR. Proper information sharing would lead to better threat and vulnerability characterization, not only in the early development and authorization stages, but throughout the entire system life cycle. Additionally, information sharing would also help improve RMF Step Six, continuous monitoring, by utilizing threat intelligence efficiently and adapting security control and risk mitigations as threats evolve.

5. Resources

The SECNAV Cybersecurity Readiness Review (2019) discusses resources as a governance tool to achieve cybersecurity resiliency, focusing on proper resource allocations to achieve strategic objectives and to act as the first levers available for use by SECNAV to transition the Navy to becoming more information centric. The Navy needs increased resources for automation and risk aggregation and an overhaul or replacement of the current primary resource for RMF documentation, eMASS.

Each DON RMF step stands to benefit from automation. Leveraging current technologies available, Steps Two through Six could be partially automated, and Step One could be automated by creating a tool and a central repository that would ingest system categorizations and deliver recommended controls. Current technologies limit manual processes, but automation would assist in reducing the overall duration of the process. Already, NIST SP 800-53 lists hundreds of security controls, those are increasing, and only a subset of controls is applicable to each system. No cyber security expert is realistically expected to memorize the applicability and benefits of each; automation could help. Although the RMF KS contains baseline controls, it is not automated to assist the ISSE in the selection of required overlays and tailoring of systems. Automating here could prove especially valuable in optimizing the effectiveness, cost, and relevance of controls in the ever-changing system environments and threat settings. Absent these tools and capabilities, the Navy will likely find it impossible to obtain and maintain true cyber security within the ever-changing cyber threat environment and applicable mitigation measures.

Several efforts exist to automate and accelerate the RMF process, such as C2C24, RAISED, and SECDEVOPS, but, unfortunately, many are disjointed and also present areas of vulnerabilities within critical mission systems as a result of U.S. Joint Staff Interoperability Requirements (SECNAV, 2019, p. 29). Interoperable systems should possess the ability to directly exchange services satisfactorily between them or their users (Joint Chiefs of Staff [JCS], 2019). For example, RAISED is designed to automate and accelerate the RMF process within C2C24; however, it only works on CANES networks. Although these efforts lean forward, security is still often an afterthought despite the RMF process guide stating the opposite. To increase efficiency, security and risk management

must be incorporated from Step Zero: Prepare. Because reducing risks to an acceptable level remains the focal point of the RMF, automation would help provide a means to collect data metrics, observe progress, improve decisions, and yield a measurable process.

Risk aggregation is an important step within the DON RMF necessary to fully assess and manage risks of our systems. According to the DON RMF Process Guide, the SCA is charged with quantifying and aggregating risk. However, both eMASS and the RMF KS are inadequate at facilitating the aggregation of risk via vulnerability chaining. Unfortunately, due to the current structure of the DON RMF and the lack of automated tools used to assist the SCA, risk aggregation is convoluted and, as a result, is often more an afterthought than the focal point. The SCA must utilize products produced throughout the other steps of the RMF process such as the SP, SAR, eMASS entries, and the KS to aggregate risk, resulting in tedious work influenced by personnel subjectivity. As long as the DON RMF remains qualitative in nature, it will not facilitate a way to aggregate risk, further substantiating the need for a standardized, more quantitative DON RMF.

When something is broken, there are two options: fix or replace. eMASS either needs replacement by a system specifically designed to support the current RMF or repair of its several broken pieces. Originating from DIACAP, eMASS was created to support Information Assurance (IA) program management and has since been adapted and updated to support RMF, resulting in significant limitations. NIST SP 800-55 (2008) shows that the data from CCRI, CSICP, Penetration tests, and more is supposed to be used to support measurements and risk assessments, but eMASS is incapable of ingesting, processing, and correlating findings from these reports. In addition to the reports specifically called out in NIST SP 800-55, the RMF needs a tool capable of ingesting other forms of data that possess the potential to affect risk management decisions such as threat intelligence and tailored recommendations, as well as additional measurements such as culture, awareness, and training.

F. CHAPTER CONCLUSION

Through step-by-step examination of the RMF process and its governing documents, our research demonstrates the need for a standardized approach to assessing

and measuring risk within the DON RMF. Key findings of the SECNAV Cybersecurity Readiness Review published in 2019 regarding issues within the DON's culture, people, structure, processes, and resources also demonstrate this need. Overall, because the DON has yet to establish any uniform or effective metrics to quantify risk, and the current DON CS and RMF processes remain ambiguous, the Navy does not yet have the agility needed to operate in the current cyber threat environment (SECNAV, 2019). Because the Navy has a critical need for such agility in risk assessment, Chapter V offers detailed recommendations to improve the RMF process.

V. RECOMMENDATIONS

This thesis examined risk and risk assessment in detail to offer potential improvements to the Navy's RMF process especially regarding information systems and Cyber Security. We reviewed the history of trying to ensure IA/CS, examined underpinnings of risk assessment, and explained how the A&A processes are currently designed. We also analyzed the current DON RMF processes, applying best practices and methods derived in the literature review, to assess current strengths and weaknesses. Chapter V summarizes our findings and offers recommendations—based both on our own operational experiences and this research—on how to improve the DON RMF practices. This chapter further offers a summary and three recommendations sections. As history shows, improvement of risk assessment processes is a continuous job, so we also conclude with suggestions for further research on the DON RMF or whatever replacement might follow.

A. SUMMARY

For quite a while, the DOD/DON has been playing catch up trying to address the exponentially increasing challenges that cyber security presents. While the current RMF approach improves upon its predecessors, it, too, needs an overhaul. Derived from NIST and DOD directives, the DON's RMF process blindly inherits the ambiguity necessary for writing policy for a polyglot of organizations and fails to tailor the RMF to specific Navy organizational need and practices, despite the specific instruction to do so in the parent instruction. Additionally, the DON RMF is highly qualitative and lacks standardized definitions, measurements, metrics, and a risk assessment methodology. The current RMF's qualitative approach is further complicated by the bias, heuristics, groupthink, inconsistency, overconfidence, and overestimation that can ensue from subjective inputs throughout the DON RMF. The DON needs a more quantitative RMF with standardized definitions, measurements, metrics, and better training to ensure appropriate risk mitigation and to provide continuous feedback for process improvement leading to increased cybersecurity and resiliency of naval networks. The following sections highlight the specific recommendations that may help the Navy accomplish that mission.

B. RECOMMENDATIONS

As we considered translating our analysis into recommendations, one overarching thought occurred. Should the Navy simply abandon the RMF and start anew, or should the Navy try to improve the existing RMF? The first option offers the chance to start afresh with new procedures that would satisfy the current RMF's many shortcomings. A new risk management framework might well include replacing the unwieldy eMASS product, improving the ease of data collection and employment, and reinforcing a more quantitative approach, yet the RMF is young by process standards. Organizations are only now becoming adept at understanding the procedures, and, because it is a new process, most participants understand that RMF has shortcomings. We, therefore, recommend option two, stay with the current RMF and make continuous improvements. As noted above, the Navy's RMF still does not contain the tailoring dictated by the NIST and DOD instructions, so that the implementation of such continuous improvements needs to be a Navy focus. Additionally, our recommendations come in three groups, described in the following sections:

- Get the basics right
- Improve each RMF step
- Address Navy-wide challenges that limit RMF

1. Get the Basics Right

We offer "getting the basics right" recommendations in three parts: standardizing definitions, establishing standard, and more quantitative, measurement methodology, and migrating from qualitative to quantitative risk matrices.

Standardize Definitions

To establish standardized definitions, we recommend DON CIO convene a working group; at a minimum, the Navy needs to define *risk*, *uncertainty*, *measurement*, and *cyber resiliency.* Additionally, the working group could consider defining other terms that result from our further recommendations. We specifically recommend that the group consider implementing the following definitions as a basis for the standards:

- Risk

  Consider and modify as needed this research's definition of risk, derived from similar definitions used by Hubbard and Seiersen, (2016) and NIST "a measure of uncertainty to which the likelihood and/or impact of a threat, potential circumstance, or event may be expressed in quantitative terms."

- Uncertainty

  Adopt Hubbard and Seiersen (2016, p. 29)'s definition of uncertainty as "the lack of complete certainty, that is, the existence of more than one possibility. The 'true' outcome/state/result/value is not known."

- Measurement

  Again, adopt the proposed definition by Hubbard and Seiersen (2016, p. 21) of measurement as "a quantitatively expressed reduction of uncertainty based on one or more observations."

- Cyber Resiliency

  Employ this research's definition of cyber resiliency, derived from similar definitions used by MITRE and NIST, as "an organization's ability to predict, absorb, recover, and pivot from adversity in the cyber realm while continuing to fight."

Establish Standard Quantitative Measurement Methodology

The DON RMF aims to reduce risk to an acceptable level, and deliberate measurements are, therefore, required. The words *performance* and *measurement*, or any reference to NIST SP-800-55, are rare within DoDI 8510.01 and the DON process guide. Additionally, no RMF documentation being used to determine the effectiveness of risk management throughout the life cycle of the system contains measurements. The absence of a standard measurement methodology and metrics contradicts the primary goal of RMF. Most successful Fortune 500 companies, contrastingly, do use and make available specific measurement metrics and technologies, and the Navy can leverage these to improve measurement within the DON RMF.

While we feel strongly that quantitative measuring is a must for a successful RMF program, we caution that organizations recognize that they "get what they measure." If DON RMF institutes and executes new measures that do not align with the overall objective of superior operational cyber resiliency, but instead become an organizational burden of little value, i.e., just another "check the box" formality, then those are the wrong measures. DON will need to frequently review the usefulness of their new quantitative measures. We recommend convening a working group to:

- Develop DON specific metrics to adequately track the current status and progress of our organization;
- Coordinate with NIST to update and tailor NIST SP 800-55 to DOD RMF requirements; and
- Expand upon this guidance by tailoring these measurements for application to the Navy's operational environments.

Foundationally, having the appropriate measurements in place would allow for the proper allocation and justification of resources and investments as needed, leading to a more effective RMF. See Table 1 for possible RMF process metrics.

Table 1. Examples of Standard Quantified Measures

Measure	Unit	Reasoning
Number of controls implemented	Count	Controls add security; therefore, more controls should equal a more secure system.
Number of controls modified or further tailored for additional requirements such as AOR, new threat intel, changes in enemy TTPs, et cetera	Count	Tracking required deviations to controls would demonstrate how the Navy is adjusting to threat intel and counter evolving adversary TTPs.
Percentage of deficiencies identified in penetration testing	Number of deficiencies/ Number of test elements	This is one type of test that determines if the controls are effective.
Average time to deploy system patches	Hours to deploy patches/ Number of patches	Keeping systems updated on patches is an important factor in maintaining security.
Average time it takes a new threat intel bulletin to be shared throughout all Navy Organizations (PM, ISO, SYSCOMs, SCAs, et cetera)	Hours to acknowledge new threat/New intel reported	Threat intel sharing is an important factor in maintaining security across the Navy.

Migrate from Qualitative to Quantitative Risk Matrices

If the use of risk matrices must remain within the RMF process, careful considerations must be taken to account for as many of the matrices' qualitative aspects as possible since risk matrices are particularly subject to bias, heuristics, and groupthink. We recommend the DON convene a working group of risk experts with quantitative risk assessment backgrounds to:

- Investigate the specific usage of risk matrices within DON RMF.

As a starting point, the quantitative working group should:

- Leverage Louis Anthony Cox's study, What's Wrong with Risk Matrices (2008, pp. 501–506), where he developed three axioms and one rule to improve the quantification within risk matrices: the weak consistency axiom, between-ness axiom, consistent coloring axiom, and the three-color rule.

Alternatively, the quantitative working group could research different methods that could replace or add value to risk matrices such as the Factor Analysis of Information Risk

(FAIR) (Hanes et al., 2017, pp. 262–266) and Intel's Threat Agent Risk Assessment (TARA) (Rosenquist, 2012) methods.

2. DON RMF Process

Deriving the DON RMF from the NIST RMF provides many advantages and allows the Navy to leverage the best practices and lessons learned from industry leaders and experts. However, our analysis demonstrates current shortfalls. The Navy should:

- Foster and develop more involved partnerships with NIST to assist with better addressing more of the unique challenges faced by the Navy within Cybersecurity and RMF;
- Revise internal DON processes to evaluate, adapt, and implement changes or updates made to NIST RMF policies that are applicable to the DON RMF process; and also
- Work to streamline those processes.

An updated, standardized, better defined, measurable, and more quantitative RMF would address many issues within the RMF process. The following sections identify specific improvement recommendations for each DON RMF process step.

Step Zero – Prepare

- Conduct a complete review of RMF documentation to ensure that it includes the most recent NIST updates and changes such as the prepare step outlined in NIST SP 800-37 (Revision 2).

Step One – Categorize System

The categorization step is inherently subjective, and its subjectivity impacts the remaining steps of the RMF process and budget allocation; therefore, brokering agreement on categorization between stakeholders and SMEs adds too much time to the process. Correct categorization of the system is crucial since it impacts all remaining RMF steps,

and errors here result in expensive rework. We recommend that one of the highest priorities in updating the DON RMF should be to:

- Reduce the subjectivity of system categorization in one of the two following ways.

Rather than developing categorizations based on what the acquisition community believes, mission stakeholders, including operators, maintainers, operational testers, and independent information security experts must be incorporated early in the RMF process to assist with the proper system categorization. Alternatively, consider making the categorization process independent of the stakeholders, run by a standing team of operational and RMF risk experts.

Step Two – Select Security Controls

Implementing automation in Step Two would improve the selection of security controls and help increase speed in the RMF process. Although the RMF KS is a good resource to consult in the selection of system security control and overlays, it is tedious, subjective, and inconsistent. A new or converted automated tool must be capable of data analysis to measure consistency among system categorization, security control selection, and ISSE history. Since the ISSE selects and implements the controls, a record of their history would reduce deviation and assist in removing subjective input errors. ISSE history would allow ISSEs to examine similar systems and look to improve security control selection and tailoring based on their previous experience combined with new threat intelligence. We recommend the Navy improve RMF Step Two in four ways:

- Review and update the RMF KS, expanding beyond baseline controls to assist in overlay selection and tailoring to increase KS's capacity to ingest observable data relative to security controls;

- Develop an automated tool capable of utilizing a centralized repository for all security controls of systems and their effectiveness, which must assist in the selection of overlays and tailoring based off observable measurements of system;.

- Require the NQV to be present during controls selection rather than being consulted on an "as needed" basis; and
- Establish performance metrics for the consistency of each ISSE, such as tracking each ISSE's control and overlay selection history, tailoring decisions compared to the average, and/or amount of experience and number of programs.

See Table 2 for examples of ISSE performance metrics.

Table 2. Examples of ISSE Performance Metrics for Consistency

Measure	Unit	Reasoning
Percentage of similar or identical systems ISSE has implemented	Number of same controls, overlays, and tailoring/ Number of similar or identical systems	Security controls implemented by ISSEs should be consistent among similar or identical systems.
Number of controls and overlays selected per system found to be Non-Compliant, result in a vulnerability, or are improperly implemented	Count	Frequent feedback improves consistency and reduces overconfidence. This observation would assist in determining whether an ISSE is improving or declining based on system type.
Percentage of concurrence from other ISSEs working on similar or identical systems.	Number of ISSEs concurring with control selection, overlays, tailoring/Number of similar or identical systems two or more ISSEs have worked on	Consistency amongst ISSEs is required across Navy organizations. Deviations require additional examination of system controls. This measure can also assist in identifying overconfidence.
Years of ISSE specific experience and # systems under purview	Number of Years specific ISSE experience/Number of systems under purview	Combined with measures above, how many years on average does it take to be a consistent and accurate ISSE? How many systems might that take on average?

Step Three – Implement Security Controls

Our analysis demonstrated organizational changes that deviate from written guidance throughout the RMF, which must be addressed to reduce unnecessary risk due to confusing or contradicting governing documents. We offer two specific recommendations to better implement security controls:

- When updating the DON RMF process guide, examine internal processes to remove organizational deviations that exist across current DON process; and
- Review the current testing tools used within the DON, standardize testing tools across the DON, and expand, as necessary. Systems that cannot be scanned via DON-approved tools incur further risk and should not be permitted to proceed in the RMF process.

Step Four – Assess Security Controls

Subjective interpretation of the assessment of security control results form the basis from which SCAs recommend system authorization to the AO. As a crucial RMF process step, Step Four would greatly benefit from a standardized organizational quantitative risk methodology. The fact that validators, SCAs, and SCA liaisons have no specialized risk training yet are charged with assessing risk is alarming. Misinterpretation and inappropriate risk assignment within Step Four provides skewed data to the AO, who then has the potential to authorize a system with great risk to the Navy. To avoid such serious errors, we recommend that the Navy:

- Require all validators, SCAs, and SCA liaisons to complete specialized risk training;
- Identify how often vulnerabilities discovered in Step Four were not identified in Step Three, which may indicate either that testing tools are not sufficient or that training is lacking for test tool operations; and
- Establish metrics that track the NQV, SCA, and SCA Liaison's consistency in risk assessments throughout their career, such as amount of experience and

training as SCAs, vulnerabilities discovered in Step Four missed in Step Three, and bias statistics based on individual SCAs versus the average SCA performance.

See Table 3 for examples of possible performance metrics for personnel consistency.

Table 3. Examples of NQV, SCA Liaison, SCA Performance Metrics for Consistency

Measure	Unit	Reasoning
Percentage of similar or identical systems NQV, SCA Liaison, SCA have assessed.	Number of systems assessed with different levels of risk/ Number of similar or identical systems assessed	The NQV, SCA Liaison, SCA should consistently assess similar or identical systems with the same level of risk and reasoning.
Percentage of concurrence from other NQV, SCA Liaison, SCA assessing similar or identical systems.	Number of NQVs, SCA Liaison, SCA concurring with assessment of risk /Number of similar or identical systems 2 or more NQVs, SCA Liaisons, SCAs have assessed	Consistency amongst NQVs, SCA Liaisons, SCAs is required across Navy Organizations. Deviations require additional examination of risk assessments.
Percentage of systems NQV assesses that SCA Liaison or SCA are directly involved in risk assessment from the beginning.	Number of systems NQV assesses with direct involvement from SCA Liaison or SCA in risk assessment from the beginning/Number of total systems NQV has assessed	This measurement may help determine whether the NQV needs more or less assistance in assessing risk.

Step Five – Authorize System

Until the RMF update addresses the issues related to the previous steps, AOs will continue to face significant challenges. AOs need the most relevant and accurate picture of risk to make informed decisions. The Navy can measure all AOs for consistency throughout their career and for consistency compared to other Navy AOs. See Table 4 for examples of potential AO consistency metrics.

Table 4. Examples of AO Performance Metrics for Consistency

Measure	Unit	Reasoning
Percentage of instances an AO does not concur with SCA recommendation	Number of systems AO did not concur with/Total number of systems reviewed by AO.	This measure would provide insight as to how often AOs disagree with SCA recommendation and perhaps illuminate whether AOs are actually determining risk.
Number of systems not approved by the AO	Count	This measure would provide insight as to how often specific AOs deny authorizations.
Number of systems an AO authorized to operate with a waiver or extension	Count	This measure will help track the shift from compliance to effectiveness by being able to observe decreasing instances of waivers and waiver extensions specific to each AO.
Number of systems approved by an AO which have been exploited	Count	This measure would allow specific tracking of systems approved by an AO which have been exploited and may require further examination.
Percentage of similar or identical systems AO has authorized with varying risk levels	Number of similar or identical systems authorized with varying risk levels /Number of similar or identical systems reviewed	Consistency amongst AOs is required across Navy organizations. Deviations require additional examination.
Percentage of concurrence from other AOs reviewing similar or identical systems.	Number of AOs concurring with assessment of risk and decision/Number of similar or identical systems 2 or more AOs have assessed	Consistency amongst AOs is required across Navy organizations. Deviations require additional examination.

We recommend the Navy establish additional metrics to document system waivers granted by AOs for those risks that do not meet the RMF requirements. We offer the following specific additional recommendations regarding AOs:

- Establish minimum training, education, and experience requirements to serve as AO, including AO PQS;

- Consider a pay bonus for qualifying as an AO;
- Recruit, train, or appoint specialized risk experts for consultation by AOs;
- Establish metrics for tracking AO consistency; and
- Create a tool that tracks the number of waivers and extensions for waivers issued, including notes on the AO's justification for the waiver and proposed operational workarounds.

One model for such a tool is the Navy surface warfare combat system capabilities and limitations secure website, which lists the results of every system test, identifies trouble reports, and offers tactical workarounds. This site is responsible for increasing warfighting resilience.

Step Six – Monitor Security Controls

DON RMF Step Six lacks clear definition and specific guidance and is actually not implemented in but rather after the DON RMF process, yet this continuous monitoring is intended for reassessments as needed throughout a system's life cycle. Implementing Step Six requires the integration of a system's history including the original RMF package, inspections such as CSICP, CCORI, and penetration testing, and threat intelligence reports in one virtual location so that the SCA, AO, and senior leadership can execute continuous monitoring. To do so more effectively, we recommend that the Navy:

- Define and standardize continuous monitoring and continuous reassessment;
- Leverage automation, analysis, and data collection into a virtual RMF workplace available to the SCA and AO; and
- Develop a dashboard that is accessible to the commanding officers and system operators that allows for near-real time network status, analysis, and monitoring.

3. Navy-Wide Challenges that Limit RMF

The SECNAV report highlights issues with culture, people, structure, process, and resource challenges in CS today. Our analysis applied those issues directly to the DON RMF. Recommendations for decreasing the impact of those issues follow.

Culture

The DON needs a cultural shift from episodic compliance to continuous cyber warfighting effectiveness. Doing so will require changes in accountability, incentive, invention, and measured competency.

- Accountability must be present from the most senior officials down to new accessions, and everyone must be held accountable for their actions.

- Incentivizing our commanders and people should involve recognition and reward for exceeding the expectations of cyber security and resiliency, not for achieving the bare minimum.

- Creative thinking, through innovation and adaptation of new technologies, will incentivize our Sailors to do better things rather than being stuck in the mindset of "this is how it has always been done."

- Measured competency applies to both commanders as well as subordinate personnel. At a minimum, if a commander is evaluated on their proficiency with damage control, maintenance, and warfighting, for example, there should also be a proficiency score associated with cyber security with the same expected of subordinate personnel on both the officer and enlisted side. The days of clicking through the same old cyber awareness course on NKO need to disappear. We need a more robust approach to measuring the understanding of cybersecurity within our force. This could be accomplished via observation, yearly proficiency testing, email tests to address spear phishing and other tactics.

- More robust training should be included in every Navy accession source and should be required for promotion and advancement.

Leading cultural change is easier said than done. In order to affect a change throughout the fleet, a sound plan must be developed, implemented, and continuously evaluated for improvement. Numerous authors have developed methodologies for leading changes throughout organizations. One approach recommended by the authors is John Kotter's methodology for Leading Change and his Eight Steps to Accelerate Change in Your Organization (Kotter, 2018).

People

People are the DON's greatest asset and sometimes its biggest liability. Military, DOD civilian, and contractors alike need to be properly trained on RMF, cyber security, and resiliency, and on how such impacts the organization as a whole. The SECNAV Cybersecurity Readiness Review (2019) proposed the Identify-Recruit-Train-Sustain-Retain model to assist with the many people-issues present throughout the fleet. Based on this study, and our own operational warfighting experience, we recommend the Navy:

- Consult with the best and brightest risk experts in the private sector. These bright minds should be leveraged by the DON CIO appointed working groups identified in these recommendations. Use them to assist in developing a standardized quantitative methodology and approach to risk assessments. Have them develop a series of standardized DOD and DON certifications that prepare and support our risk assessors to execute their duties;

- Develop a formally accredited and rigorous process for DON risk certification. The DON needs individuals who are thoroughly familiar with assessing risk and can assist in addressing the needed improvements for RMF. Develop risk certifications for RMF personnel that meet similar standards for risk assessors in the insurance, financial, and safety fields;

- Research expanding the CTN ratings to employ them at the most tactical and operational levels within naval networks, since most ITs are not trained to

monitor the network and the RMF or to fight the ship through a cyber-attack; and,

- Alternatively, overhaul IT training to include RMF, continuous monitoring, and fighting of naval networks since there are far more ITs than CTNs.

IT training was overhauled several years ago with spectacular results, using software from Acuitus (S. Miller, personal communication, May 6, 2020). However, that overhaul was abandoned because of shortsighted cost avoidance by NETC (S. Miller, personal communication, May 6, 2020). Consider reimplementing that approach for IT "A" school.

Organizational Structure

Thorough examination of the current DON RMF process and RMF leadership structure identified a convoluted and fragmented organizational structure that inhibits effective execution of RMF. The DON CIO should be a full-time and dedicated individual—with proper authority, responsibility, and accountability—who reports directly to SECNAV. We also recommend that the DON appoint a Chief Information Risk Officer (CIRO) who is subordinate to the CIO and is the delegated authority of the DON RMF. The CIRO should ideally come from private industry and have ample experience dealing with risk in various organizations. The CIRO should lead a consolidated revamping of DON RMF in conjunction with the CIO.

Process

The existence of so many stove-piped organizational structures, as well as classification and interoperability complications throughout the fleet, other Services, and Agencies plague the DOD and inhibit proper sharing of intelligence and information. That plague also negatively impacts the DON RMF. Better intelligence and information sharing would lead to improved threat and vulnerability characterization in the RMF's early stages of development, throughout the process, and during continuous monitoring through system life cycles.

The DON must decide on a definitive way forward to address the lack of intelligence sharing whether it is JIE, JEDI, or something else. Until accurate and real time

threat assessments and capabilities are incorporated into the RMF process, risk estimates will ultimately be merely a snapshot in time at best. Systems are operated in several AORs, and security controls implemented need to be tailored appropriately according to the TTPs of each respective AOR's adversaries. Absent this capability, RMF will remain insufficient.

RMF Resources

Finally, and crucially, if truly interested in cyber warfare proficiency and resiliency:

- *The Navy must increase resources devoted to this warfighting goal.*

For example, each step within the DON RMF process stands to benefit from automation. Leveraging current available technologies, Steps Two through Six can be partially automated and Step One could be automated upon creation of a centralized repository of system categorizations with ingest capabilities. We also recommend:

- Because risk aggregation remains a convoluted and resource intensive task, modify or replace the RMF KS and eMASS tools so that they:
    - Possess the capability to ingest reports produced throughout the RMF process, other third-party inspections like CSICP, CCORI, and penetration testing, as well as threat intelligence affecting cyber systems and
    - Support risk aggregation in support of cyber resiliency.

C. AREAS FOR FURTHER STUDY

Many questions remained unanswered for migrating from compliance to operational effectiveness. Future researchers might consider RMF studies in the following areas:

- Establishing and using CYBER Figure of Merit;
- Using the Cyber Range in support of system, unit, and force training;
- Exploring how a platform becomes cyber resilient;

- Researching how to migrate from a firewall-based security at NOCs to individual-based security;
- Because true cyber resiliency implies a highly educated crew, identifying ways to accelerate learning so that cyber defenders can become experts;
- Studying whether risk matrices and their usage within the Navy have improved risk-based decisions or reduced risk;
- Studying alternate risk methodologies to supplement or replace the use of risk matrices such as the FAIR and TARA methods;
- Designing a new KS capable of ingesting observable data relative to security controls; and
- Determining how to establish calibration training for key players in the DON RMF process.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Adams, A. (2015, October 16). How group dynamics affect decisions. Stanford News. https://news.stanford.edu/features/2015/decisions/group-dynamics.html

Allahverdyan, A., & Galstyan, A. (2014). Opinion dynamics with confirmation bias. *PLOS ONE*, *9*(7), e99557. https://doi.org/10.1371/journal.pone.0099557

Albert, D., Bedek, M., Huszar, L., & Nussbaumer, A. (2017). Effekt von clustering illusion (cognitive bias) bei benutzung einer "visual analytics" Umgebung. forschungsdaten zur studie 2017; Effect of clustering illusion during the interaction with a visual analytics environment. Research data of the 2017 study. https://doi.org/10.5160/PSYCHDATA.DHAT17EF10

Angner, E. (2006). Economists as experts: Overconfidence in theory and practice. *Journal of Economic Methodology*, *13*(1), 1–24. https://doi.org/10.1080/13501780600566271

Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, *193*(5), 2–6.

Asch, S. (1956). Studies of independence and conformity: 1.A minority of one against a unanimous majority. *Psychological Monographs*, *70*(9), 1–70. https://doi.org/10.1037/h0093718

Ball, D., & Watt, J. (2013). Further thoughts on the utility of risk matrices. *Risk Analysis*, *33*(11), 2068–2078. https://doi.org/10.1111/risa.12057

Bang, D., & Frith, C. (2017). Making better decisions in groups. *Royal Society Open Science*, *4*(8), 170193. https://doi.org/10.1098/rsos.170193

Barrett, D. (2017). United States Navy Risk Management Framework process guide Version 2.0. Washington, DC: Department of Navy.

Bazerman, M. H., & Moore, D. A. (2013). *Judgment in managerial decision making* (8th ed.). Wiley.

Begh, R., Munafò, M.R., Shiffman, S, Ferguson, S.G., Nichols, L., Mohammed, M.A., Holder, R.L., Sutton, S. & Aveyard, P. (2013). Attentional bias retraining in cigarette smokers attempting smoking cessation (ARTS): Study protocol for a double blind randomised controlled trial. *BMC Public Health 13*(1). https://doi.org/10.1186/1471-2458-13-1176

Bhatia, S. (2015). The power of the representativeness heuristic. *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 232–238). https://www.sas.upenn.edu/~bhatiasu/Bhatia%202015%20CogSci%20PP.pdf

Brookins, P., & Ryvkin, D. (2014). An experimental study of bidding in contests of incomplete information. *Experimental Economics, 17*(2), 245–261. http://doi.org/http://dx.doi.org/10.1007/s10683-013-9365-9

Budescu, David, Broomell, Stephen & Por, Han-Hui. (2009). Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. *Psychological science. 20*. 299–308. https://doi.org/10.1111/j.1467-9280.2009.02284.x.

Cherry, K. (2020, January 23). How the attentional bias influences the decisions we Make. Very Well Mind. https://www.verywellmind.com/what-is-an-attentional-bias-2795027

Clore, G., & Huntsinger, J. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*(9), 393–399. http://doi.org/10.1016(J.tics.2007.08.005

Colwell, L. H. (2005). Cognitive heuristics in the context of legal decision making. *American Journal of Forensic Psychology, 23*(2), 17–41. https://psycnet.apa.org/record/2005-06088-002

Committee on National Security Systems Glossary No. 4009, Committee on National Security Systems Glossary No. 4009 (2015). Retrieved from https://rmf.org/wp-content/uploads/2017/10/CNSSI-4009.pdf

Cox, L. A. Jr., Babayev, D., & Huber, W. (2005). Some limitations of qualitative risk rating systems. *Risk Analysis*, *25*(3), 651– 662. https://doi.org/10.1111/j.1539-6924.2005.00615.x

Cox, L. A. Jr. (2008). What's wrong with risk matrices? *Risk Analysis*, *28*(2), 497–512. https://doi.org/10.1111/j.1539-6924.2008.01030.x

Davis, P. K., Kulick, J., & Egner, M. (2005). Implications of modern decision science for military decision-support systems. RAND. https://ebookcentral.proquest.com

Dawes, R. (1994) *House of cards: Psychology and psychotherapy built on myth*. Free Press.

Department of Defense. (n.d.). Model for assessing residual risk level for non-compliant security controls. RMF Knowledge Service. Retrieved March 19, 2020, from, https://rmfks.osd.mil/rmf/RMFImplementation/AssessControls/Pages/ResidualRisk.aspx

Department of Defense. (1997, December 30). Department of Defense Information Technology Security Certification and Accreditation Process (DITSCAP). (DoDI 5200.40) Department of Defense. http://www.acqnotes.com/Attachments/DOD%20Instruction%205200.40.pdf

Department of Defense. (2017, January). Department of Defense risk, issue, and opportunity management guide for defense acquisition. Department of Defense. https://www.dau.edu/tools/Lists/DAUTools/Attachments/140/RIO-Guide-January2017.pdf

Department of Defense. (2007, March 12). DOD Information Assurance Certification and Accreditation Process (DIACAP). (DoDI 8510.01). Department of Defense. http://www.acqnotes.com/Attachments/DOD%20Instruction%208510.01.pdf

Dowd, K. W., Petrocelli, J. V, & Wood, M. T. (2014). Integrating information from multiple sources: A metacognitive account of self-generated and externally provided anchors. *Thinking & Reasoning, 20*(3), 315–332. https://doi.org/10.1080/13546783.2013.811442

Dvorsky, G. (2013, September 1). The 12 cognitive biases that prevent you from being rational. GIZMODO. https://io9.gizmodo.com/the-12-cognitive-biases-that-prevent-you-from-being-rat-5974468

Epley, N. (2013). Anchoring. In *Encyclopedia of the mind*. http://dx.doi.org/10.4135/9781452257044.n11

Evans, J., Newstead, S., & Byrne, R. (1993). *Human reasoning: the psychology of deduction*. Lawrence Erlbaum Associates.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 13*(1), 1–17. https://doi.org/10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S

Fischhoff, B., Slavic, P., & Lichtenstein, S. (1982). Lay foibles and expert fables in judgments about risk. *The American Statistician*, *36*(3), 240–255, DOI: 10.1080/00031305.1982.10482845

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases, 422–24. Cambridge University Press. https://apps.dtic.mil/dtic/tr/fulltext/u2/a099435.pdf

Frick, W. (2015, February 2). What research tells us about making accurate predictions. *Harvard Business Review*. https://hbr.org/2015/02/what-research-tells-us-about-making-accurate-predictions

Gardner, D. (2010). Future babble: Why expert predictions fail-and why we believe them anyway. McClelland & Stewart.

Gilovich, T. (1991). How we know what isn't so. The fallibility of human reason in everyday life. Free Press.

Gilovich, T., Vallone, R. & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*(3), 295–314. https://doi.org/10.1016/0010-0285(85)90010-6

Goldberg, L. R. (1968). Simple models or simple processes? Some research on clinical judgments. *American Psychologist, 23*(7), 483–496. doi: http://dx.doi.org.libproxy.nps.edu/10.1037/h0026206

Greene, M. (2018). Insurance. In *Encyclopaedia Britannica*. https://www.britannica.com/topic/insurance

Greer, L., Caruso, H., & Jehn, K. (2011). The bigger they are, the harder they fall: Linking team power, team conflict, and performance. *Organizational Behavior and Human Decision Processes*, *116*(1), 116–128. https://doi.org/10.1016/j.obhdp.2011.03.005

Hanes, D., Salgueiro, G., Grossetete, P., Barton, R., & Henry, J. (2017). *IOT Fundamentals: Networking technologies, protocols, and use cases for the internet of things.* Cisco Press.

Haselton, M. G.; Nettle, D. & Andrews, P. W. (2005). *The evolution of cognitive bias*. In D. M. Buss (Ed.), The handbook of evolutionary psychology, pp. 724–746. John Wiley & Sons.

Heilbronner, S. R., Hayden, B. Y., & Platt, M. L. (2010). Neuroeconomics of risk-sensitive decision making. In G. J. Madden & W. K. Bickel (Eds.), Impulsivity: The behavioral and neurological science of discounting (p. 159–187). American Psychological Association. https://doi.org/10.1037/12069-006

Herrmann, A. (2013). The quantitative estimation of IT-related risk probabilities. *Risk Analysis 33*(8) 1510–1531. https://doi.org/10.1111/risa.12001

Heuer, R. (1999). *Psychology of intelligence analysis*. Center for the study of intelligence, Central Intelligence Agency.

Hoelzl, E., & Rustichini, A. (2005). Overconfident: Do you put your money on it? *Economic Journal, 115*(503), 305–318. http://www.blackwell-synergy.com/doi/abs/10.1111/j.1468-0297.2005.00990.x

Hubbard, D. (2009). *The failure of risk management: Why its broken and how to fix it.* John Wiley & Sons.

Hubbard, D. (2014). *How to measure anything: Finding the value of intangibles in business.* John Wiley & Sons.

Hubbard, D. & Seiersen, R. (2016). *How to measure anything in cybersecurity risk*. John Wiley & Sons.

Ibtida, R., & Pamungkas, B., (2018). The design of a risk-based performance audit program for court-fee management for the comptroller of Supreme Court of Indonesia. Proceedings of the Advances in Economics, Business and Management Research (AEBMR), 6th International Accounting Conference (IAC 2017), 55, 201–206. https://doi.org/10.2991/iac-17.2018.36

International Organization for Standardization. (2009). ISO 73:2009 (Risk management-vocabulary). Retrieved December 17, 2019, from https://www.iso.org/obp/ui#iso:std:iso:13824:ed-1:v1:en:term:3.8

ISACA. (n.d.). Glossary. (n.d.). Retrieved December 17, 2019, from https://www.isaca.org/Pages/Glossary.aspx?tid=1784&char=R.

Janis, I. (1972). Victims of groupthink: *A psychological study of foreign-policy decisions and fiascoes.* Houghton Mifflin.

Joffe, H. (2003). Risk: From perception to social representation. *The British Journal of Social Psychology, 42(*55). http://search.proquest.com/docview/219172806/

Joint Chiefs of Staff. (2019) Joint communications system (JP 6-0). Retrieved from https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp6_0ch1.pdf?ver=2019-10-15-172254-827

Juslin, P. (2013). Availability heuristic. In H. Pashler (Ed.), *Encyclopedia of the mind, 1*, 103–104. SAGE Publications. doi: 10.4135/9781452257044.n39

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430–454. https://doi.org/10.1016/0010-0285(72)90016-3

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341–350.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition, 11*(2), 123–141. https://doi.org/10.1016/0010-0277(82)90022-1

Kahneman, D. (2003). A perspective on judgment and choice. *American Psychologist, 58*(9), 697–720. http://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D. (2011). *Thinking fast and slow*. Farrar, Straus and Giroux.

Katopol, P. (2015). Groupthink: Group dynamics and the decision-making process. *Library Leadership & Management, 30*(1). https://journals.tdl.org/llm/index.php/llm/article/view/7172

Keil, M., Depledge, G. and Rai, A. (2007), Escalation: the role of problem recognition and cognitive bias, *Decision Sciences, 38*(3), 391–421. https://doi.org/10.1111/j.1540-5915.2007.00164.x

Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes, 39*(1), 98–114. https://doi.org/10.1016/0749-5978(87)90047-1

Keren G. (1991). Calibration and probability judgments: conceptual and methodological issues. *Acta Psychologica, 77*(3), 217–273. https://doi.org/10.1016/0001-6918(91)90036-Y

Klipstein, M. (2017). Quantifying risk for decentralized offensive cyber operations. [Doctoral dissertation, Naval Postgraduate School]. NPS Archive: Calhoun. http://hdl.handle.net/10945/54879

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for Confidence. *Journal of Experimental Psychology: Human learning and memory 6*(2), 107–18. http://iipdm.haifa.ac.il/images/publications/Asher_Koriat/1980-Koriat-Lichtenstein-Fischhoff-JEPHLM.pdf

Kotter, J. (2018). *8 steps to accelerate change in your organization*. Kotter. Retrieved 2020, from https://www.kotterinc.com/wp-content/uploads/2019/04/8-Steps-eBook-Kotter-2018.pdf

Krawczyk, M. W., Rachubik, J. (2019). The representativeness heuristic and the choice of lottery tickets: A field experiment. *Judgment and Decision Making, 14*(1), 51–57. http://journal.sjdm.org/18/18318/jdm18318.pdf

Levine, E. (2012). Improving risk matrices: the advantages of logarithmically scaled axes. *Journal of Risk Research*, *15*(2), 209–222. https://doi.org/10.1080/13669877.2011.634514

Lichtenstein, S., Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behaviour and Human Decision Processes*, 20(2) 159– 183. https://doi.org/10.1016/0030-5073(77)90001-0

Lichtenstein S, Fischhoff B, Phillips L., D. (1981). Calibration of probabilities: The state of the art 1980 (Report No. PTR-1092-81-6). Office of Naval Research. http://www.ccnss.org/ccn_2014/materials/pdf/sigman/callibration_probabilities_lichtenstein_fischoff_philips.pdf

Markovits, H. & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition 17*, 11–17. https://doi.org/10.3758/BF03199552

Markowski A. S., Mannan M. S. (2008). Fuzzy risk matrix. *Journal of Hazardous Materials, 159*(1), 152–157. https://doi.org/10.1016/j.jhazmat.2008.03.055

Mauck, J., & Pashley, C. (2016, January-March). Changing the DOD cyberspace culture. DON CIO. https://www.doncio.navy.mil/mobile/ContentView.aspx?ID=7387&TypeID=21

Mcbride, M., Fidler, F., & Burgman, M. (2012). Evaluating the accuracy and calibration of expert predictions under uncertainty: predicting the outcomes of ecological research. *Diversity and Distributions*, *18*(8), 782–794. https://doi.org/10.1111/j.1472-4642.2012.00884.x

Miller, N. R., Kiriakou, C. M., & Hilton, P. (2015). Navy authorizing official and security control assessor risk management framework process guide. Fort Meade, MD: U.S. Fleet Cyber Command. This document is For Official Use Only.

Monti, S., & Carenini, G. (2000). Dealing with the expert inconsistency in probability elicitation. *Knowledge and Data Engineering, IEEE Transactions On*, *12*(4), 499–508. doi: 10.1109/69.868903

Moore, D., Swift, S., Minster, A., Mellers, B., Ungar, L., Tetlock, P., Yang, H., & Tenney, E. (2017). Confidence calibration in a multi-year geopolitical forecasting competition. *Management Science, 63*(11), 3552–3565. https://doi.org/10.1287/mnsc.2016.2525

Mothes, C. (2017). Confirmation bias. In *The SAGE encyclopedia of political behavior.* https://sk.sagepub.com/reference/the-sage-encyclopedia-of-political-behavior/i2441.xml

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies und shortcomings of social judgment.* Prentice-Hall.

Office of the Chief of Naval Operations. (2018, July 18). U.S. NAVY Cyber Security Program (OPNAVINST 5239.1D). Department of the Navy. https://www.secnav.navy.mil/doni/Directives/05000%20General%20Management%20Security%20and%20Safety%20Services/05-200%20Management%20Program%20and%20Techniques%20Services/5239.1D.pdf

Olson, C. L. (1976). Some apparent violations of the representativeness heuristic in human judgment. *Journal of Experimental Psychology: Human Perception and Performance, 2*(4), 599–608. https://doi.org/10.1037/0096-1523.2.4.599

Office of the Deputy Chief of Naval Operations for Information Dominance (OPNAV N2/N6). (2014, October-December) Task Force Cyber Awakening OPNAV N2/N6 Takes the lead. *CHIPS.* https://www.doncio.navy.mil/chips/ArticleDetails.aspx?ID=5394

Orr, D., & Guthrie, C. (2006). Anchoring, information, expertise, and negotiation: New insights from meta-analysis. *Ohio State Journal on Dispute Resolution, 21*(3), 597–628. http://hdl.handle.net/1811/77238

Oskarsson, A., Van Boven, L., Mcclelland, G., & Hastie, R. (2009). What's next? Judging sequences of binary events. *Psychological Bulletin*, *135*(2), 262–285. https://doi.org/10.1037/a0014821

Peace, C. (2017). The risk matrix: uncertain results? *Policy and Practice in Health and Safety*, *15*(2), 131–144. https://doi.org/10.1080/14773996.2017.1348571

Pfleeger, S., & Caputo, D. (2012), Leveraging behavioral science to mitigate cyber security risk, *Computers & Security, 31*(4), pp. 597–611. https://doi.org/10.1016/j.cose.2011.12.010

Phillips-Wren, G., Power, D., & Mora, M. (2019) Cognitive bias, decision styles, and risk attitudes in decision making and DSS, *Journal of Decision Systems, 28*(2), 63–66. DOI: 10.1080/12460125.2019.1646509

Prims, J., & Moore, D., (2017). Overconfidence over the lifespan. *Judgment and Decision Making*, *12*(1), 29–41. http://journal.sjdm.org/15/151024/jdm151024.pdf

Ramler, R. & Felderer, M., (2013). Experiences from an initial study on risk probability estimation based on expert opinion. *Proceedings of the Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement*, 93–97. https://doi.org/10.1109/IWSM-Mensura.2013.23

Roggi, O., & Altman, E. (2013). *Managing and measuring of risk: Emerging global standards and regulations after the financial crisis* (Vol. 5). World Scientific Publishing Co Pte Ltd.

Romona, S. (2011). Advantages and disadvantages of quantitative and qualitative information risk approaches. *Chinese Business Review, 10*(12), 1106–1110. https://doi.org/10.17265/1537-1506/2011.12.002

Rosenquist, M. (2012, August 20). Top 10 questions for the threat agent risk assessment (TARA) methodology. *Intel IT peer network.* https://itpeernetwork.intel.com/top-10-questions-for-the-threat-agent-risk-assessment-tara-methodology/#gs.6snawj

Ross, T., Wu B., Kreinovich, V., (2000). Optimal elimination of inconsistency in expert knowledge: Formulation of the problem fast algorithms. *Proceedings of the International Conference on Intelligent Technologies*, 450–458. https://scholarworks.utep.edu/cs_techrep/492/

Rot, A. (2008). IT risk assessment: Quantitative and qualitative approach. *Proceedings of the World Congress on Engineering and Computer Science*, 1073–1078. http://www.iaeng.org/publication/WCECS2008/WCECS2008_pp1073-1078.pdf

Savage, S. (2002). The flaw of averages. *Harvard Business Review*, *80*(11), 20–21. http://search.proquest.com/docview/227832736/

Secretary of the Navy. (2019). Secretary of the Navy cybersecurity readiness review. https://www.navy.mil/strategic/CyberSecurityReview.pdf

Silver, N. (2012). *The signal and the noise: Why so many predictions fail--but some don't.* Penguin Group.

Slovic, P., Finucane, M., Peters, E., & Macgregor, D. (2007). The affect heuristic. *European Journal of Operational Research, 177*(3), 1333–1352. https://doi.org/10.1016/j.ejor.2005.04.006

Sternberg, R., Leighton, J. (2004). *The nature of reasoning*. The Press Syndicate of the University of Cambridge.

Tetlock P., (2005). *Expert Political Judgment*. Princeton University Press.

Thomas, P., Bratvold, R., & Bickel, J. (2013). The risk of using risk matrices. Proceeding of the *Society of Petroleum Engineers - SPE Annual Technical Conference and Exhibition, ATCE 2013, 6, 56–66*. https://doi.org/10.2118/166269-MS

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293–315. https://doi.org/10.1037/0033-295X.90.4.293

U.S. Dept. of Commerce, National Institute of Standards and Technology. Performance measurement guide for information security, Revision 1 (2008). Gaithersburg, MD. Retrieved from https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-55r1.pdf

U.S. Department of Commerce, National Institute of Standards and Technology. Guide for conducting risk assessments, (2012). Gaithersburg, MD. Retrieved from https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf

U.S. Department of Commerce, National Institute of Standards and Technology. Risk Management Framework for information systems and organizations: A system life cycle approach for security and privacy (2018). Gaithersburg, MD. Retrieved from https://csrc.nist.gov/publications/detail/sp/800-37/rev-2/final

U.S. Department of Commerce, National Institute of Standards and Technology. Managing information security risk: organization, mission, and information system view, Managing information security risk: organization, mission, and information system view (2011). Gaithersburg, MD. Retrieved from https://csrc.nist.gov/publications/detail/sp/800-39/final

U.S. Department of Commerce, National Institute of Standards and Technology. Security and privacy controls for federal information systems and organization (2015). Gaithersburg, MD. Retrieved from https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r4.pdf

Valladares, J. (2013). Effectiveness of the Department of Defense information assurance accreditation process. [Master's thesis, Army War College]. Retrieved from http://www.dtic.mil/docs/citations/ADA590269

Vazzano Ltd. (n.d.). History of marine cargo insurance. Retrieved December 15, 2019, from http://www.cargoins.com/history

Wall, K. (2011). The trouble with risk matrices. [Working paper]. Naval Postgraduate School http://hdl.handle.net/10945/32570

Wickham, P.A. (2003), "The representativeness heuristic in judgments involving entrepreneurial success and failure," *Management Decision, 41*(2), 156–167. https://doi.org/10.1108/00251740310457605

Williams, P. & Steward, T. (2007, September-October). DOD's Information Assurance Certification & Accreditation Process, *Defense Acquisition Technology and Logistics 36*(5), 12–13.

Wouter Botzen, W., Kunreuther, H., & Michel-Kerjan, E. (2015). Divergence between individual perceptions and objective indicators of tail risks: Evidence from floodplain residents in New York City. *Judgment and Decision Making*, *10*(4), 365–385. http://journal.sjdm.org/15/15415/jdm15415.html

Yudkowsky, E. (2008). Cognitive biases potentially affecting judgment of global risks. In Bostrom, N. & Cirkovic, M. (Eds.), *Global Catastrophic Risks,* 91–119. Oxford University Press.

Zajonc, R. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2). 151–175. https://doi.org/10.1037/0003-066X.35.2.151

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California