NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

PSYCHOMETRIC ANALYSIS OF THE SERVICEMEMBER EVALUATION TOOL

by

Joseph S. Felix

March 2018

Thesis Advisor:	Brennan Cox
Co-Advisor:	Andrew Anglemyer
Second Reader:	Nita Shattuck

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704–0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** March 2018	**3. REPORT TYPE AND DATES COVERED** Master's thesis	
4. TITLE AND SUBTITLE PSYCHOMETRIC ANALYSIS OF THE SERVICEMEMBER EVALUATION TOOL			**5. FUNDING NUMBERS**
6. AUTHOR(S) Joseph S. Felix			
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. IRB number NPS.2017.0022-AM04-EP4&7-A .			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

The U.S. Navy is concerned with the psychological resilience of its sailors. The Navy has developed multiple programs as part of the 21st Century Sailor and Marine initiative to promote resilient behaviors. In order to determine the effectiveness of these programs, the Navy must have a valid tool to assess psychological resilience. The Naval Center for Combat and Operational Stress Control developed the Servicemember Evaluation Tool (SET) to assess its internal resilience-training program. This thesis examines the use of the SET in an operational setting by collecting data from five naval vessels on deployment. The goals of this thesis are to evaluate the psychometric properties of the SET on this sample population. Furthermore, this thesis identifies potential improvements to the SET to make it more efficient for operational use.

14. SUBJECT TERMS resilience, psychometrics, Servicemember Evaluation Tool			**15. NUMBER OF PAGES** 95
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** 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.

PSYCHOMETRIC ANALYSIS OF THE SERVICEMEMBER EVALUATION TOOL

Joseph S. Felix
Lieutenant Commander, United States Navy
B.S., Case Western Reserve University, 2004

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 2018

Approved by: Brennan Cox, Ph.D.
Thesis Advisor

Andrew Anglemyer, Ph.D.
Co-Advisor

Nita Shattuck, Ph.D.
Second Reader

Patricia A. Jacobs, Ph.D.
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The U.S. Navy is concerned with the psychological resilience of its sailors. The Navy has developed multiple programs as part of the 21st Century Sailor and Marine initiative to promote resilient behaviors. In order to determine the effectiveness of these programs, the Navy must have a valid tool to assess psychological resilience. The Naval Center for Combat and Operational Stress Control developed the Servicemember Evaluation Tool (SET) to assess its internal resilience-training program. This thesis examines the use of the SET in an operational setting by collecting data from five naval vessels on deployment. The goals of this thesis are to evaluate the psychometric properties of the SET on this sample population. Furthermore, this thesis identifies potential improvements to the SET to make it more efficient for operational use.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION 1
- A. BACKGROUND 1
- B. OBJECTIVES 2
  - 1. Measuring the Psychometric Properties of SET 2
  - 2. Determining Potential Improvements to SET 4
- C. SCOPE 5

II. LITERATURE REVIEW 7
- A. DEFINING RESILIENCE 7
- B. RESILIENCE IN THE MILITARY 8
- C. REVIEW OF LITERATURE ON SET COMPONENTS 9

III. METHODS 15
- A. PARTICIPANTS 15
- B. MEASURES 15
- C. ANALYSES 16
  - 1. Sample Demographics 16
  - 2. Individual Item-Level Analysis 17
  - 3. Scale Level Analysis 17
  - 4. Factor Analysis 17
  - 5. Battery-Wide Analysis of SET 18

IV. RESULTS 21
- A. SAMPLE DEMOGRAPHICS 21
- B. INDIVIDUAL ITEM LEVEL ANALYSIS 23
  - 1. Epworth Sleepiness Scale 23
  - 2. Insomnia Severity Index 24
  - 3. Pittsburgh Sleep Quality Index 25
  - 4. Connor-Davidson Resilience Scale 26
  - 5. Response to Stressful Experiences Scale 27
  - 6. Brief COPE 27
  - 7. New General Self-Efficacy Scale 28
  - 8. Brief Self-Control Scale 29
  - 9. Life Orientation Test-Revised 30
  - 10. Thought Control Questionnaire 31
  - 11. Cognitive Emotion Regulation Questionnaire 32
  - 12. Unit Social Support Scale 32

13. NCCOSC Morale Questions 33
14. Satisfaction with Life Scale 34
15. Short Form-36 General Health Question 34
16. Perceived Stress Scale.....35
17. Primary Care PTSD Screen.....36
18. Patient Health Questionnaire.....36
C. SCALE LEVEL ANALYSIS 37
D. FACTOR ANALYSIS 38
1. Epworth Sleepiness Scale 38
2. Insomnia Severity Index.....39
3. Pittsburgh Sleep Quality Index.....40
4. Connor-Davidson Resilience Scale 41
5. Response to Stressful Experiences Scale.....42
6. Brief COPE.....42
7. New General Self-Efficacy Scale.....43
8. Brief Self-Control Scale.....44
9. Life Orientation Test-Revised.....45
10. Thought Control Questionnaire 46
11. Cognitive Emotion Regulation Questionnaire 47
12. Unit Social Support Scale.....47
13. NCCOSC Morale Questions 48
14. Satisfaction with Life Scale 49
15. Short Form–36 General Health Question.....49
16. Perceived Stress Scale.....49
17. Primary Care PTSD Screen.....50
18. Patient Health Questionnaire.....50
E. BATTERY-WIDE ANALYSIS OF SET 51

V. CONCLUSIONS AND RECOMMENDATIONS.....53
A. RECOMMENDATIONS TO SHORTEN THE SET BATTERY 53
B. CONCLUSIONS 55
C. LIMITATIONS.....56
D. RECOMMENDATIONS FOR FUTURE WORK.....57

APPENDIX. SET BATTERY 59

LIST OF REFERENCES.....69

INITIAL DISTRIBUTION LIST 75

LIST OF TABLES

Table 1. Number of Participants from each Ship 21

Table 2. Demographic Data Comparison 22

Table 3. Health-Related Behaviors of SET Participants 23

Table 4. Epworth Sleepiness Scale Item Distribution 24

Table 5. Insomnia Severity Index Item Distribution 25

Table 6. Pittsburgh Sleep Quality Index Component Distribution 26

Table 7. Connor-Davidson Resilience Scale Item Distribution 26

Table 8. Response to Stressful Experiences Item Distribution 27

Table 9. Brief COPE Item Distribution 28

Table 10. New General Self-Efficacy Scale Item Distribution 29

Table 11. Brief Self-Control Scale Item Distribution 30

Table 12. Life Orientation Test-Revised Item Distribution 31

Table 13. Thought Control Questionnaire Item Distribution 31

Table 14. Cognitive Emotion Regulation Questionnaire Item Distribution 32

Table 15. Unit Social Support Scale Item Distribution 33

Table 16. NCCOSC Morale Item Distribution 33

Table 17. Table N. Satisfaction with Life Scale Item Distribution 34

Table 18. SF-36 General Health Questionnaire Item Distribution 35

Table 19. Perceived Stress Scale Item Distribution 35

Table 20. Primary Care PTSD Screen Item Distribution 36

Table 21. Patient Health Questionnaire Item Distribution 37

Table 22. Comparison of Cronbach's Alpha with Published Results 37

Table 23. Factor Loadings for Epworth Sleepiness Scale 39

Table 24. Factor Loadings for Insomnia Severity Index 40
Table 25. Factor Loadings for Pittsburgh Sleep Quality Index Components 41
Table 26. Factor Loadings for Connor-Davidson Resilience Scale.....41
Table 27. Factor Loading for Response to Stressful Experiences Scale.....42
Table 28. Factor Loadings for Brief COPE 43
Table 29. Factor Loadings for New General Self-Efficacy Scale.....44
Table 30. Factor Loadings for the Brief Self-Control Scale 45
Table 31. Factor Loadings for Life Orientation Test-Revised.....46
Table 32. Factor Loadings for the Thought Control Questionnaire.....46
Table 33. Factor Loadings for Cognitive Emotion Regulation Questionnaire 47
Table 34. Factor Loadings for the Unit Social Support Scale 48
Table 35. Factor loadings for the NCCOSC Morale Questions.....48
Table 36. Factor Loadings for the Satisfaction with Life Scale.....49
Table 37. Factor Loadings for Perceived Stress Scale.....50
Table 38. Factor Loadings for Primary Care PTSD Screen.....50
Table 39. Factor Loadings for the Patient Health Questionnaire.....51
Table 40. Correlation of Individual Scale Scores 52
Table 41. Recommended Scales to Retain or Eliminate from the SET 55

LIST OF ACRONYMS AND ABBREVIATIONS

BC	Brief COPE
BSS	Brief Self-control Scale
CAF	Comprehensive Airman Fitness program
CAPS	Clinician Administered Scale for PTSD
CD-RISC	Connor-Davidson Resilience Scale
CFA	confirmatory factor analysis
CERQ	Cognitive Emotion Regulation Questionnaire
CJCS	Chairman of the Joint Chiefs of Staff
CSF2	Comprehensive Soldier and Family Fitness program
DRRI	Deployment Risk and Resiliency Inventory
DoD	Department of Defense
ESS	Epworth Sleepiness Scale
EFA	exploratory factor analysis
IRB	institutional review board
ISI	Insomnia Severity Index
LOT-R	Life Orientation Test-Revised
MOS	Medical Outcomes Study
NCCOSC	Naval Center for Combat and Operational Stress Control
NGSES	New General Self-Efficacy Scale
NPS	Naval Postgraduate School
ODASD (MC&FP)	Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy
PC-PTSD	Primary Care PTSD screen
PCL	PTSD Symptom Checklist
PHQ	Patient Health Questionnaire
PSQI	Pittsburgh Sleep Quality Index
PSS	Perceived Stress Scale
PTSD	post-traumatic stress disorder
RSES	Response to Stressful Experiences Scale
SECNAV	Secretary of the Navy

SET	Servicemember Evaluation Tool
SF-36	Short Form 36 questionnaire
SWLS	Satisfaction with Life Scale
TCQ	Thought Control Questionnaire
USS	Unit Social Support Scale

EXECUTIVE SUMMARY

The U.S. Navy is more than ships, aircraft, and other platforms; it is people. The Navy invests a considerable portion of its budget in recruiting, training, and retaining the best people in order to operate its platforms in support of the national military strategy. The Navy seeks to protect its investment in people through programs such as the 21st Century Sailor and Marine initiative. This program aims "to maximize Sailor and Marine personal readiness, to maintain the resilience of the force, and to hone the most combat effective force possible" (Secretary of the Navy [SECNAV], 2013, para. 1).

This thesis focuses on the measurement of resilience within active duty sailors in the U.S. Navy. The Naval Center for Combat and Operational Stress Control (NCCOSC) developed the Servicemember Evaluation Tool (SET) as a means to assess their internal resilience training programs. The overall goal of this study is to evaluate the practical utility of the SET as an operational measure of sailor resilience. The SET is a battery comprised of 18 individual scales, each of which assesses a subdomain or facet related to the larger resilience construct. These aspects include sleep, optimism, positive reappraisal, morale, unit support, self-efficacy, self-control, satisfaction, and overall health. Additionally, the SET contains scales that seek to measure resilience in general. The first objective of this thesis is to measure the psychometric properties of the SET to determine if the scales are reliable and valid. The second objective is to determine what improvements to make to the SET to make it more efficient for operational use.

The participants in this study were part of a larger research project on human performance that looked at the differences between sailors on ships that employed a circadian watch rotation versus ships that used a non-circadian rotation. Researchers from the Naval Postgraduate School (NPS) recruited sailors for voluntary participation in the study while embarked on the USS *Chafee* (DDG-90), USS *Pinckney* (DDG-91), USS *Kidd* (DDG-100), and USS *Princeton* (CG-59) for approximately two weeks at the beginning of their deployments in June 2017. Additionally, some sailors from the USS *Vella Gulf* (CG-72) completed the SET during their deployment and mailed their responses back to NPS.

I conducted item, scale, and battery-wide analyses on the SET. The item level analysis looked at each of the 134 survey items individually. Following guidance by DeVellis (2003), I evaluated each item for grammar, readability, and understandability. I examined all item response distributions and descriptive statistics, including mean, median, standard deviation, and skewness. At the scale level, I evaluated the internal consistency of the scales by calculating Cronbach's alpha. I also performed factor analyses on the scales to identify the number of latent variables measured by each scale. For all scale level analyses, I compared results with those found in the published literature, when available. For the battery-wide analysis, I compared the correlations between scores on each of the scales in the SET battery. Based on the results of all levels of analysis, I made recommendations for which scales to retain and which scales to eliminate from the SET. The purpose of these recommendations was to reduce the time needed to complete the SET battery while maintaining a valid and reliable tool for assessing resilience.

The results showed that all scales were generally psychometrically sound. The individual items tended to have normal response distributions. The internal consistency estimates and factor structure of most scales matched those in previous research. I attribute these results to NCCOSC adopting established scales instead of creating entirely new measures for the SET. Even though some scales were modified slightly, they performed well in this study with this population.

The result of the battery-wide analysis was a recommendation to shorten the SET from 18 scales down to nine. This analysis reduced the number of items by over 45%. Table 1 gives the recommendations to retain or eliminate for each scale in the SET battery. Among the remaining scales, all but the Brief Cope (BC) and the Patient Health Questionnaire (PHQ) have a five-point Likert response. For consistency, I would change the BC and PHQ from a four-point scale to a five-point scale. I would also review the instructions for each scale to make sure the time span they reference is consistent. For example, some scales ask participants to consider the past month while other scales are based on the previous week or two weeks.

Table 1. Recommendation to Retain or Eliminate Scale from the SET

Scales to Retain	Scales to Eliminate
Insomnia Severity Index	Epworth Sleepiness Scale
Connor-Davidson Resilience Scale	Pittsburgh Sleep Quality Index
Brief COPE	Response to Stressful Experiences Scale
New General Self-Efficacy Scale	Brief Self-Control Scale
Life Orientation Test-Revised	Thought Control Questionnaire
Cognitive Emotion Regulation Questionnaire	NCCOSC Morale Questions
Unit Social Support Scale	Satisfaction with Life Scale
Perceived Stress Scale	SF-36 General Health Question
Patient Health Questionnaire	Primary Care PTSD Screen

This study lays the groundwork for many potential avenues of research on the SET and resilience. Results from the SET could be analyzed with data gathered as part of the larger study on human performance to compare the resilience of crewmembers on circadian versus non-circadian watch rotations. Researchers could also analyze the data with regard to other breakout groups such as age, rank, gender, ship, or department. Furthermore, researchers could use the shortened version of the SET battery in future data collections. They could collect data from sailors in other environments such as aviation, submarines, amphibious vessels, or shore duty. All of these studies would further validate the usefulness of SET in assessing the resilience of Sailors and Marines.

References

DeVellis, R. F. (2003). *Scale development: theory and applications* (Vols. 26, Applied Social Research Methods). Thousand Oaks, CA: SAGE.

Secretary of the Navy. (2013). *Department of the Navy 21st Century Sailor and Marine initiative* (SECNAVINST 5300.40*).* Washington, DC: Department of the Navy.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank my advisors, Dr. Brennan Cox and Dr. Andy Anglemyer, for their support and guidance in completing this thesis. I would like to thank Dr. Nita Shattuck and the staff and students on the NPS Crew Endurance Research team who went underway to collect data for this study. I am extremely grateful for Josh Strubel and his team of STEM interns who completed a majority of the data entry. I am appreciative of Dr. Eileen Delaney of the Naval Center for Combat and Operational Stress Control for providing valuable background information on the development of the SET. Finally, I would like to thank my wife, Martha, who supported me throughout this process and took care of everything while I was busy writing this thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

The U.S. Navy is more than ships, aircraft, and other platforms; it is people. The Navy invests a considerable portion of its budget in recruiting, training, and retaining the best people in order to operate its platforms in support of the national military strategy. Serving in the military often involves being in dangerous and stressful environments. Lives are at risk every time an aircraft takes off, a submarine submerges, or a ship gets underway. Deployments away from friends, family, and other support may add to the stress experienced by sailors. In addition to physical fitness, sailors must maintain psychological fitness.

The Navy seeks to protect its investment in people through programs such as the 21st Century Sailor and Marine initiative. This program aims "to maximize Sailor and Marine personal readiness, to maintain the resilience of the force, and to hone the most combat effective force possible" (Secretary of the Navy [SECNAV], 2013, para. 1). This initiative includes a number of programs related to alcohol abuse, suicide prevention, safety, physical fitness, and others. The effectiveness of some of these programs can be assessed through clear and objective indicators, such as the number of safety mishaps or the number of failures on a physical fitness test. Measuring the resilience of the force is a more difficult task.

This thesis focuses on the measurement of resilience within active duty sailors in the U.S. Navy. Specifically, this thesis examines the Servicemember Evaluation Tool (SET) developed by the Naval Center for Combat and Operational Stress Control (NCCOSC). The Navy currently has several programs that seek to improve the resilience and readiness of its sailors. However, "the Navy and Marine Corps do not have a mechanism to systematically and efficiently assess the psychological readiness of its fleet and forces, diminishing the ability to make data-driven decisions about unit- and force-level wellbeing and resource allocation" (NCCOSC, 2016, para. 2).

NCCOSC initially created the SET as a means to assess the performance of their internal resilience training programs. However, NCCOSC has recognized the opportunity to use SET to assess the psychological resilience of the U.S. Navy fleet and forces. NCCOSC has experimented with using different combinations of scales in the SET battery depending on the situation. They would like to identify a core set of resilience measures that can be used in all situations (E. Delaney, personal communication, September 12, 2017). In order to achieve this end, the SET must demonstrate that it is a reliable, valid, and practical instrument.

B. OBJECTIVES

The overall goal of this study is to evaluate the practical utility of the SET as an operational measure of sailor resilience. The two objectives of this thesis are in support of this goal. The first objective is to measure the psychometric properties of the SET to determine if the scales are reliable and valid. The second objective is to determine what improvements to make to the SET to make it more efficient for operational use.

1. Measuring the Psychometric Properties of SET

Because resilience is a psychological construct, it is not directly observable. Researchers can only infer resilience indirectly through the observation of behaviors associated with the resilience construct. The measurement of psychological constructs such as resilience is not as straightforward as measuring physical concepts such as height or weight. Psychological constructs are often measured by surveys or questionnaires known as scales. In the same way a thermometer must be calibrated to assess temperature, psychological scales must also be evaluated and refined to ensure they produce reliable and valid measurements; this process is known as psychometrics. According to Litwin (2003), "psychometrics provides survey researchers with a way to quantify the precision of the measurement of qualitative concepts" (p. 1).

Measurement precision is critical to ensuring a survey is measuring the construct it was designed to assess. In this case, the SET is only a useful tool if it consistently and accurately measures the resilience of service members. Although there are several ways to

measure reliability and validity (Litwin, 2003), those most pertinent to this study include internal consistency reliability, content validity, and construct validity.

"Internal consistency reliability, as the name implies, is concerned with the homogeneity of the items within a scale" (DeVellis, 2003, p. 27). Researchers typically report internal consistency using the Cronbach's alpha statistic, and support use of .07 as an acceptable cutoff value for basic research (e.g., Nunnally, 1978; Peterson, 1994). Common practice whenever using an existing scale on a new population is to determine whether the internal consistency estimate is comparable to values found in the published literature. Scales that demonstrate acceptable levels of internal consistency across populations are more reliable than those that exhibit acceptable internal consistency for some populations and unacceptable internal consistency for others.

Content validity describes the extent to which the elements of a survey are representative of the particular construct they are designed to measure (Litwin, 2003). As a battery of resilience measures, the SET would demonstrate content validity if its items and scales were shown to effectively represent the full range of elements under the resilience umbrella. NCCOSC developed the SET in large part based on the findings of a RAND report (Meredith et al., 2011) on resilience in the U.S. military. This report identifies 20 unique resilience factors that correspond to four general categories: individual, family, unit, and community. Of these, the SET contains measures that correspond to the individual and unit categories only. In order to demonstrate content validity, the SET should demonstrate adequate coverage of these factors.

Litwin (2003) describes construct validity as "a measure of how meaningful the scale or survey instrument is when in practical use" (p. 41). Construct validity is often comprised of convergent and divergent validity. The scores on a particular scale have convergent validity if they correlate to other scales that measure the same construct; they have divergent validity if they do not correlate closely with scales that measure a different construct (Litwin, 2003). Demonstrating convergent validity in the SET battery is the first step in establishing construct validity. This is important for determining if the SET battery is a meaningful measure of resilience for service members.

Factor analysis is another type of psychometric analysis. The objective of a factor analysis is "to help an investigator determine how many latent variables underlie a set of items" (DeVellis, 2003, p. 103). There are two reasons for identifying the latent variables. First, examining the latent variables helps with assessing content validity because it shows which factors related to resilience the particular scales cover. Second, a scale is a candidate for elimination from the SET battery if another scale already measures the same latent variable. In many cases, previous researchers have already completed a factor analysis on a particular scale so my goal is to confirm these results.

In order for the SET to be considered "psychometrically sound," each scale needs to demonstrate acceptable levels of reliability and validity. The use of internal consistency, content validity, factor analysis, and convergent validity are all appropriate for analyzing the psychometric properties of this measure as needed to achieve the first objective of this study.

2. Determining Potential Improvements to SET

Even if a scale is psychometrically sound, there are other considerations to determine if it has practical utility. For instance, NCCOSC recognizes that the length of the SET in its current form may not be practical for all uses (E. Delaney, personal communication, September 12, 2017). Since the SET battery is composed of multiple scales, there may be overlap among the scales in the aspects of resilience it measures. In addition, the SET battery contains scales that measure clinical symptoms and health behaviors that may be related to resilience (NCCOSC, 2016). One potential improvement is to eliminate redundancies in the SET.

The second objective of this thesis is to use the psychometric analysis of the SET as the basis for reducing the overall length of the SET. The number of items can make a difference in the effectiveness of the scale. Too few items can reduce the reliability of the scale, and too many items can be burdensome to those filling out the survey (DeVellis, 2003). A shortened version of the SET would increase is applicability in operational settings. Using a data-driven approach, this thesis aims to reduce the number of items and scales on the SET without sacrificing the overall reliability and validity of the battery.

C. SCOPE

This thesis is one part of a larger research project examining human performance considerations on five U.S. Navy ships over the course of a deployment and was approved by the Naval Postgraduate School (NPS) Institutional Review Board (IRB; protocol #2017.0022-AM04-EP4&7-A). The larger project included research on sleep patterns, watch rotations, reaction times, individual performance, and command performance, and was longitudinal by design. Data for this thesis were collected during the initial data collection period of the larger study. Although the larger study assessed differences in human performance between ships operating with a circadian watch rotation compared to ships operating with a non-circadian rotation, this thesis is exploratory in nature and focuses mainly on the SET measure itself.

THIS PAGE INTENTIONALLY LEFT BLANK

II. LITERATURE REVIEW

A. DEFINING RESILIENCE

"Resilience" comes from the Latin *resiliens* meaning to rebound or recoil ("Etymonline," n.d.). The term was initially used in the metallurgical sense of a material returning to its original size or shape after being strained (Fletcher & Sarkar, 2013). Over time, however, and similar to the concept of stress, resilience has been adapted for use in psychological contexts to describe how individuals respond to challenging events.

Resilience is a complex construct without a universally accepted definition. For instance, Meredith et al. (2011) reported 122 separate definitions of psychological resilience found throughout the published literature. As summarized by Fletcher and Sarkar (2013), some definitions view resilience as a process (e.g., Masten, Best, & Garmezy, 1990; Luthar, Cicchetti, & Becker, 2000), while others view it as an ability (e.g., Bonanno, 2004) or a set of behavioral tendencies (Agaibi & Wilson, 2005). Yet despite these differences, most definitions of psychological resilience include the concepts of adversity and positive adaptation (Meredith et al., 2011).

Adversity concerns the challenges of life, ranging from everyday hassles (e.g., traffic, lost keys, workplace demands, and interpersonal relationships) to major traumatic events (e.g., serious injury, natural disasters, bankruptcy, and death). Of note, adversity can also be associated with positive life events; the birth of a child, for instance, while generally viewed as a positive life event, entails many stressors intermixed with excitement.

Individuals differ in their response to adversity. For examples, some people view adversity as a challenge, and rise to the occasion, while others view adverse events more as a threat, and become defensive or avoidant. Richardson, Neiger, Jensen, and Kumpfer (1990) suggest that, when stressful events cause a disruption in a person's life, they may return to a higher level of resilience, the same level of resilience, a lower level of resilience, or a dysfunctional state with suicidal behavior or drug and alcohol abuse. This positive response to adversity is key to the understanding of resilience.

According to the Chairman of the Joint Chiefs of Staff (CJCS), resilience is the "ability to withstand, recover, and grow in the face of stressors and changing demands" (CJCS, 2011, Glossary). This definition assumes that resilience is an ability, but it is an ability military personnel can improve through training. This definition also includes the basic elements of adversity and positive adaptation. For these reasons, I adopted this definition of resilience for the purpose of this thesis.

B. RESILIENCE IN THE MILITARY

Due to the nature of military service, service members frequently experience both physical and psychological stressors. Just as physical fitness is necessary to survive the physical hardships of combat, psychological fitness is necessary for military personnel to handle the mental stresses of their duty. Psychological resilience is highly valued by the Department of Defense (DoD), as evidenced by research which has identified 23 separate programs aimed at improving the psychological resilience of military personnel (Meredith et al., 2011). One of the conclusions drawn from this RAND study was that "policy to define roles, responsibilities, and broad guidance for implementation would be extremely helpful" (Meredith et al., 2011, p. 76). This study also noted that existing military programs lack empirical evidence necessary to assess their effectiveness at building resilience.

The armed services have responded with policies to improve resilience among their Soldiers, Sailors, Marines, and Airmen. In addition to the 21st Century Sailor and Marine Initiative developed by the Navy, the Army has released the Comprehensive Soldier and Family Fitness (CSF2) program, and the Air Force has the Comprehensive Airman Fitness (CAF) program. The mission of the CSF2 program is "to increase the physical and psychological health, resilience, and performance of Soldiers, Families, and DACs [Department of the Army Civilian employees]" (Department of the Army, 2014, p. 7). For the Air Force, the "CAF strategy focuses on strengthening fitness, resilience, and readiness in Airmen, families, communities, and organizations through education, resilience building activities, and wellness support programs" (Department of the Air Force, 2014). Building a resilient force is a common thread in all three of these programs. Since all branches in

the DoD have made a goal to build resilience, it is important to be able to accurately assess the level of resilience of military personnel and to determine if this goal is being met.

C. REVIEW OF LITERATURE ON SET COMPONENTS

Consistent with the vision set forth by the Navy's 21st Century Sailor office, NCCOSC promotes the psychological health and resilience of Sailors and Marines. In support of this mission, NCCOSC developed the SET, a battery of 18 scales designed to assess content areas associated with psychological resilience. These content areas include sleep, optimism, positive reappraisal, morale, unit support, self-efficacy, self-control, satisfaction, and overall health. Additionally, the SET contains scales that seek to measure resilience in general. Details on the development and intended use of each of the SET scales are provided below.

The first three scales in the SET battery are measures related to sleep, and include the Epworth Sleepiness Scale (ESS) (Johns, 1991), the Insomnia Severity Index (ISI) (Bastien, Vallières, & Morin, 2001), and the Pittsburgh Sleep Quality Index (PSQI) (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989). The ESS contains eight items and has a four-point Likert response for a score range of 0 to 24. The ESS assesses daytime sleepiness. The ISI has seven items and a five-point Likert response for a score range of 0 to 28. The ISI measures symptoms of insomnia. The PSQI assesses sleep history using 19 total questions; four items are fill-in-the-blank and 15 items have a four-point Likert response. The scoring algorithm calculates seven component scores which add together to form a global PSQI score ranging from 0 to 21. The final item in the PSQI is to be answered by a bed partner or roommate. This item does not affect the scoring algorithm and was not included in the SET. These scales were initially developed for patients with sleep disorders; however, they have been used in research on military personnel (e.g., Shattuck & Matsangas, 2016; Bryan, 2011; Babson, Blonigen, Boden, Drescher, & Bonn-Miller, 2012).

The Connor-Davidson Resilience Scale (CD-RISC) was developed as a means to assess resilience in general (Connor & Davidson, 2003). The original scale included 25 items with five-point Likert responses. The authors tested this scale on both the general

population and patients with psychiatric disorders such as general anxiety disorder and PTSD. Campbell-Sills and Stein (2007) conducted further psychometric analysis using undergraduate students. Their research reduced the CD-RISC from 25 items to 10 and from a five-factor model to a single factor: resilience. NCCOSC adopted this 10-item CD-RISC for the SET.

The Response to Stressful Experiences Scale (RSES) was designed to complement other measures of resilience such as CD-RISC (Johnson et al., 2011). The original 22-item version of RSES has a five-point Likert response, was validated using a sample of military personnel, and resulted in a five-factor model. These five factors were meaning-making and restoration, active coping, cognitive flexibility, spirituality, and self-efficacy (Johnson et al., 2011). Like the CD-RISC, further psychometric evaluation reduced this scale to a unidimensional model containing only four items (De la Rosa, Webb-Murphy, & Johnston, 2016). These four items did not include the spirituality domain, but they did address the other four factors from the original version of RSES. This four-item RSES maintained high levels of internal consistency and was used in the SET.

The next scale in the SET is the Brief COPE. Carver (1997) introduced the 28-item Brief COPE as an update to the original 60-item COPE inventory (Carver, Scheier, Weintraub, 1989). Instead of 15 subscales with four items each, this new scale consists of two items for each of 14 subscales related to various coping techniques, such as humor, planning, and denial. The responses for these items are on a four-point Likert scale ranging from “I haven’t been doing this at all” to “I’ve been doing this a lot.” While Carver (1997) analyzed the Brief COPE based on surveys of Hurricane Andrew survivors, Rice and Liu (2016) used the Brief COPE as part of a study of resilience in active duty and veterans of the U.S. military. This study grouped the 14 coping strategies into emotion-focused coping, problem-focused coping, and dysfunctional coping strategies. The version of the Brief COPE used in the SET uses 18 items covering nine of the 14 coping strategies with six items from each of the three groups.

The New General Self-Efficacy Scale (NGSES) seeks to capture the construct of general self-efficacy (Chen, Gully, & Eden, 2001). Both general self-efficacy and specific self-efficacy refer to one’s belief in their own competency in a situation, but the scope for

specific self-efficacy is narrowed to a specific situation whereas general self-efficacy applies to a wider range of situations. The NGSES includes eight items with five-point Likert responses. Chen et al. (2001) have demonstrated that the NGSES has a high level of internal consistency and reliability.

Another concept that plays a role in resilience is self-control. Tangney, Baumeister, and Boone (2004) developed the Brief Self-Control Scale (BSS). This scale uses 13 different statements and gives respondents five options ranging from "Not like me at all" to "Very much like me." Self-control is applicable to several domains including achievement and task performance, impulse control, adjustment, and interpersonal relationships (Tangney et al., 2004). The application to adjustment indicates why the BSS was included in the SET.

The Life Orientation Test (LOT) is an instrument designed to measure dispositional optimism (Scheier & Carver, 1985). The original LOT was a 12-item scale with a five-point Likert response. Four of these 12 items are positively worded, another four items are negatively worded (reversed scored), and the remaining four items are filler items that are not included in the scoring. Scheier, Carver, and Bridges (1994) conducted further research with a large sample of undergraduate students with the goal of improving the predictive validity of the LOT. This produced a revised Life Orientation Test (LOT-R) that included only 10 items. A principal component factor analysis on the six non-filler questions produced one factor with an eigenvalue greater than 1.0. Scheier et al. (1994) also conducted a confirmatory factor analysis to test a one-factor and a two-factor model. As with Scheier and Carver (1985), they noted that in the two-factor model, the positively worded items loaded to one factor and the negatively worded items loaded on the other factor. The four filler questions for the LOT-R are not used in the SET.

The next two scales on the SET measure positive reappraisal. Garnefski and Kraaij (2007) define positive reappraisal as "thoughts of creating a positive meaning to the event in terms of personal growth" (p. 142). The Thought Control Questionnaire (TCQ) measures to what extent a person uses one of five strategies for controlling unpleasant or unwanted thoughts (Wells & Davies, 1994). These strategies are distraction, social control, worry, punishment, and reappraisal. The TCQ contains 30 items and uses a four-point Likert

response option. The SET uses the six items from the TCQ that correspond to the reappraisal strategy. Wells and Davies note that reappraisal is a way to cope with unwanted thoughts by thinking logically about them. The Cognitive Emotion Regulation Questionnaire (CERQ; Garnefski & Kraaij, 2007) also assesses positive reappraisal. The full CERQ includes nine subscales using 36 items with a five-point response scale. The four items in the positive reappraisal subscale are included in the SET.

The Deployment Risk and Resiliency Inventory (DRRI) was developed to assess a range of resilience factors that affect deployed military personnel and veterans who have previously deployed (King, King, Vogt, Knight, & Samper, 2006). The revised version, DRRI-2, contains 17 separate scales based on research with veterans of the wars in Afghanistan and Iraq (Vogt et al., 2013). The Unit Social Support (USS) scale measures the effect that the unit has on an individual's resilience during deployment. The USS consists of 12 items with a five-point Likert response. It is the only section of the DRRI-2 that is included in the SET. However, NCCOSC also included a section in the SET that asks participants to rate their own level of morale and the morale in their command on a five-point Likert response scale.

The next scale in the SET is the Satisfaction with Life Scale (SWLS; Diener, Emmons, Larson, & Griffin, 1985). In this case, satisfaction is based on one's own judgment of their situation in comparison to what they think it should be. Numerous researchers have validated the SWLS with a wide variety of sample populations including college students, elderly Americans, military wives, and male prison inmates (for a summary, see Pavot & Diener, 1993). The SWLS consists of five items and is the only scale in the SET battery that uses a seven-point Likert response scale.

The Medical Outcomes Study (MOS) 36-Item Short-Form Health Survey (SF-36) was developed to capture patients' own judgment of the outcomes of their medical care (Ware & Sherbourne, 1992). The entire SF-36 assesses eight different areas of health concerns. The SET only includes the first question of the SF-36, which asks the respondent to rate their health as either excellent, very good, good, fair, or poor.

The Perceived Stress Scale is a 14-item scale with a five-point Likert response that was designed to assess perceived stress (Cohen, Kamarck, & Mermelstein, 1983). Cohen and Williamson later updated the PSS with a 10-item version and a four-item version (1988). They found that the PSS-10 performed as well if not better than the PSS-14. The PSS-4 was slightly less reliable but was still appropriate for use when brevity is desired. This four-item version of the PSS is used in the SET.

The next scale in the SET battery is a measure of Post-Traumatic Stress Disorder (PTSD) symptoms. Prins et al. (2003) developed a four-item scale with a yes or no response, known as the primary care PTSD (PC-PTSD), to quickly screen patients for PTSD symptoms. The authors compared the PC-PTSD to other assessments such as PTSD Symptom Checklist (PCL) and the Clinician Administered Scale for PTSD (CAPS). The PC-PTSD did not perform as well as the PCL, but it has the advantage of being only four items long versus 17 for the PCL. Any patient with a score of three or four on the PC-PTSD should be further evaluated using the CAPS, which is the gold standard for PTSD diagnoses.

The final scale in the SET is the four-item Patient Health Questionnaire (PHQ-4; Kroenke, Spitzer, Williams, & Löwe, 2009). This scale combines two items from a depression scale with two items from an anxiety scale. The scale uses a four-point Likert response, so scores range from 0 to 12. Kroenke et al. (2009) found the PHQ-4 to be both reliable and valid for assessing depression and anxiety.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHODS

A. PARTICIPANTS

The participants in this study were part of a larger research project on human performance in relation to their watch scheduling practices (circadian versus non-circadian). Researchers from the Naval Postgraduate School (NPS) embarked on four vessels for approximately two weeks at the beginning of their deployments in June 2017. The USS *Chafee* (DDG-90), USS *Pinckney* (DDG-91), and USS *Kidd* (DDG-100) are all Flight 2A Arleigh Burke-class guided missile destroyers. The USS *Princeton* (CG-59) is a Ticonderoga class guided missile cruiser. All four ships are part of the U.S. Pacific fleet. The NPS researchers recruited sailors of all ranks for voluntary participation. Sailors had the option of participating in a sleep study, which included filling out the SET at the end of the study, or completing the SET only. This analysis includes data from both groups. Additionally, sailors from the USS *Vella Gulf* (CG-72), a cruiser in the Atlantic fleet, had the option of completing the SET during their deployment. These sailors mailed their responses back to researchers at NPS.

B. MEASURES

In addition to the battery of scales that make up the SET, participants filled out a brief questionnaire consisting of demographic and behavior related items. The demographic data included basic information, such as age and gender, as well as military information, such as rate, rank, years of service, and the number of times deployed. The remaining items identified possible behaviors that may affect the participant's sleep or overall health. The questions asked sailors to list the type and frequency of any nicotine products used, caffeinated beverages consumed, or medications taken. Another question listed possible environmental factors that could affect the participants' sleep. The final item asked about the frequency, duration, and type of exercise routine followed by the individual.

C. ANALYSES

The research team entered the responses to the demographic questions and all 18 of the scales in the SET battery into a Microsoft Excel spreadsheet. All but one of the scales used a Likert type response scale. The spreadsheet converted these responses to numerical values in accordance with the scoring rubric for each of the individual scales. The PC-PTSD scale had a binary response scale with Yes = 1 and No = 0. Some scales included reverse-scored items. I entered missing responses as a blank instead of a zero so the value would not affect the scoring rubric. If any of the items for an individual in one scale were blank, I also assigned a blank as the overall score for that scale. In the demographics section and the PSQI questionnaire, there are fill-in-the-blank response options for some questions. If the response was a range of values, I entered the average value. For example, if a participant answered "10 to 20" for the number of minutes to fall asleep in question 2 of the PSQI, the recorded value was 15. I used the statistical software program JMP Pro (Version 13.1) to conduct further analysis of the data. I conducted five types of analysis on the data to include sample demographics, item-level analysis, scale level analysis, factor analysis, and analysis of the overall SET.

1. Sample Demographics

Participants answered questions related to their demographics and personal behaviors along with their responses to the SET questionnaire. I used the demographic information to group respondents by age, rank, gender, ship, and department. I compared these data to those provided by the Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy (ODASD (MC&FP)) (2017), which reported estimates from the Defense Manpower Data Center for the fiscal year 2015, to evaluate the degree to which the SET sample was representative of the larger Navy. The personal behaviors included questions related to sleep, caffeine use, nicotine use, medication use, and exercise routine. I calculated the percentage of respondents who exhibited these behaviors and identified the most common factor that affected their sleep.

2. Individual Item-Level Analysis

The item level analysis looked at each survey item individually. In accordance with the guidance described by DeVellis (2003), I checked each item for grammar, readability, and understandability. Missing responses could indicate that participants did not understand a particular item. I examined the distribution of responses for each item and calculated descriptive statistics such as mean, median, standard deviation, and skewness. A distribution with a positive skew has a longer tail on the right side of the distribution, which implies that the median value is less than the mean. A distribution with a negative skew has a longer tail on the left side of the distribution, which implies that the median value is larger than the mean (Kim, 2013).

3. Scale Level Analysis

The scale level analysis examines the internal consistency of the entire scale, which effectively captures the correlation between each item on a given scale, represented by Cronbach's α. Many of the scales used in the SET have values for Cronbach's α published from previous research using the scale. I compared the values calculated from the SET data with the published values from previous research on each particular scale, if available.

4. Factor Analysis

Factor analysis is the third type of analysis performed on the data. The factor is the underlying latent variable that the scale is trying to measure such as the resilience, optimism, or satisfaction. Some of the scales used in the SET have already been used in previous research in which a factor analysis was performed. In these cases, I performed a confirmatory factor analysis (CFA). For the CFA, I created a model with the same number of factors as in the previous research, and I examined the model fit.

For every factor that is identified through factor analytic procedures, factor loading estimates are calculated. Factor loadings represent the correlation between the observed score for each item and the latent variable(s) derived in the factor analysis solution. When available, I compared the loadings with values from previous research to see if the same items loaded on the same factors in the SET data as they did in previous research.

For other scales, there was no documented factor analysis or the scale used in the SET was a partial scale of a previously researched scale. In these cases, I performed an exploratory factor analysis (EFA). For the EFA, I created several models with different number of factors to explore which number of factors has the best fit. For both CFA and EFA, I followed the guidance from DeVellis (2003) for choosing the appropriate number of factors. I used both the eigenvalue rule and the scree test to determine the appropriate number of factors for each scale. The eigenvalue rule says that factors with an eigenvalue of less than one account for less information than the average item and should not be included. The scree test plots the eigenvalues for each factor in order of decreasing value. Since less relative information is gained by successive factors, the entire plot has a negative slope. However, there can be a transition from a steep negative slope to a shallow negative slope known as an "elbow." The implication is that information gained from the factors after the "elbow" is insignificant compared to the information in the factors before the "elbow" (DeVellis, 2003). I looked for an "elbow" in the scree plot and took the number of factors where the transition takes place.

5. Battery-Wide Analysis of SET

Finally, I examined the SET battery as a whole. I eliminated scales that were redundant and scales that had weaker psychometric properties. Based on personal experience from administering the SET to volunteers in the study and reports from other members of the research team, the length of the SET made it burdensome to fill out for sailors that are already pressed for time. Therefore, I aimed to develop a shorter version of the SET that was still psychometrically sound for use in future research on the resilience of sailors.

In addition to the scores on the individual items, most of the scales in the SET battery produce an overall score by summing the scores on individual items (after reverse scoring some items if necessary). Exceptions include the PSQI (Buysse et al., 1989), which first calculates seven component scores which are then added together to obtain a global score, and the BC (Carver, 1997), which does not calculate an overall score. Instead, the BC sums the scores from the two items that correspond to each coping technique. Of the

nine BC coping techniques included in the SET, I grouped Planning, Active Coping, Positive Reframing, Emotional Support, Instrumental Support, and Religion into the category of positive coping techniques; Denial, Self-blame, and Behavioral Disengagement were grouped as negative coping techniques. I reversed scored the items related to negative coping techniques and summed them together. The overall BC score had a possible range of 18 to 72 with higher scores indicating better coping ability.

I put these overall scale scores into a correlation matrix to evaluate the strength and direction of the relationship between the scales. Most scales on the SET are scored such that higher scores are associated with more desirable outcomes or characteristics. However, six of the scales in the SET battery (ESS, ISI, PSQI, PSS, PC-PTSD, and PHQ) are scored such that a high score corresponds to a negative outcome such as sleepiness, stress, PTSD, or anxiety. This difference affects the direction of relationship with these six scales.

This correlation matrix was used to assess the construct validity of the SET battery. A strong correlation indicated that the two scales might be measuring the same or a similar construct, and could therefore represent a potential case for redundancy. A weak correlation indicated the two scales may be measuring unrelated constructs, and therefore may be assessing something other than resilience. According to DeVellis (2003), "there is no cutoff that defines construct validity" (p. 54). Carlson and Herdman (2012) recommend correlations above 0.70 to establish convergent validity and to avoid correlation less than 0.50. However, this recommendation is for scales that are explicitly measuring the same construct. In the SET, with the exception of CD-RISC and RSES, the scales measured factors of resilience and not resilience directly. Therefore, I expected lower levels of correlation between the individual scales. Instead of setting concrete cutoff values, I evaluated the relative strength of each correlation value to all other correlation values in the matrix.

In order to maintain content validity, I needed to ensure scales in the SET battery sufficiently encompassed the resilience domain but did not include factors outside of the domain. I established the following rules to determine which scales to keep and which scales to eliminate. First, I examined the correlations to see how the scale correlated to other scales. If correlations were relatively similar, I compared the internal consistencies

as measured by Cronbach's alpha. Finally, I qualitatively considered the wording and response options of the items in the scales. When there were multiple scales that addressed a similar construct, I chose only one of those scales. I considered the results of the entire psychometric analysis so that recommendations for inclusion or removal were not based on a single criterion. These processes were intended to improve the SET by reducing redundant scales and removing scales that are not as psychometrically sound or do not provide useful or relevant information.

IV. RESULTS

A. SAMPLE DEMOGRAPHICS

Table 1 depicts the number of subjects who volunteered to participate in the study by ship. A total of 569 participants completed components of the SET battery; however, not all participants completed all portions of the SET. Therefore, the results reported in the sections that follow are based on the valid sample for each scale, not overall participation.

Table 1. Number of Participants from each Ship

Ship	Sleep Study	SET Only	Total
USS *Chafee*	125	8	133
USS *Pinckney*	43	13	56
USS *Kidd*	98	25	123
USS *Princeton*	68	60	128
USS *Vella Gulf*	0	129	129
Subtotal	334	235	569

In addition to completing the 18 sections of the SET, participants also completed 13 questions related to demographics and behavior. Table 2 provides a comparison between the age, gender, and rank group of the participants in this study with the overall Navy population as reported by ODASD for FY2015.

Table 2. Demographic Data Comparison

	SET Dataset	FY2015 Navy-wide data
Age		
≤ 25	276 (48.8%)	137,752 (42.6%)
26 – 30	140 (24.7%)	75,578 (23.4%)
31 – 35	88 (15.5%)	48,800 (15.1%)
36 – 40	37 (6.5%)	33,088 (10.2%)
≥ 41	25 (4.4%)	28,116 (8.7%)
Gender		
Male	441 (77.9%)	264,065 (81.7%)
Female	125 (22.1%)	59,269 (18.3%)
Rank Group		
E1 – E3	71 (12.6%)	70,893 (22.0%)
E4 – E6	367 (65.2%)	168,214 (52.0%)
E7 – E9	43 (7.6%)	30,031 (9.3%)
W2 – W5	2 (0.4%)	1,649 (0.5%)
O1 – O3	74 (13.1%)	31,791 (9.8%)
O4 – O6	6 (1.1%)	20,550 (6.4%0

Notes: Not all percentages may add up to 100% due to rounding. FY2015 Navy-wide data adapted from ODASD (MC&FP) (2017).

The SET participants were similar in age, gender, and rank as the overall Navy population. By age, the SET sample has more young sailors and fewer old sailors than the overall population. By rank, the SET sample has fewer junior enlisted (E1 – E3) and fewer senior officers (O4 – O6).

The behavioral questions concerned habits that may affect a sailor's health. Table 3 is a summary of the participants' responses to these questions. Over two-thirds of all participants did not use any nicotine products and almost three-quarters followed an exercise routine. For nicotine products, cigarettes were the most commonly used source. A vast majority of the respondents used caffeinated beverages. Nearly half of participants indicated being coffee drinkers, while one-third indicated consumption of soft drinks and energy drinks. A majority selected at least one environmental factor that negatively

affected their sleep. The most common factors were temperature, noise, and not enough time to sleep.

Table 3. Health-Related Behaviors of SET Participants

Behavior	Yes	Factor that affects sleep	Yes
Use caffeinated beverages?	458 (81.1%)	Temperature	357 (63.2%)
Have an exercise routine?	407 (72.7%)	Noise	339 (60.0%)
Use nicotine products?	185 (32.7%)	Not enough time to sleep	331 (58.7%)
Take prescribed or over-the-counter medication?	112 (20.1%)	Bedding Conditions	249 (44.1%)
		Light	225 (39.8%)
		Odors	124 (21.9%)
		Motion	84 (14.9%)

B. INDIVIDUAL ITEM LEVEL ANALYSIS

For the individual item level analysis, I will discuss the items by scale in the order the scales were included in the SET. For reference, the entire SET battery is included in the appendix. Although most scales are not numbered, I added item numbers for ease of referencing specific items in the discussion. I provide item level descriptive statistics for each scale, including number of valid responses, mean, standard deviation, and skewness. Since nearly all items had kurtosis values in an acceptable range, I did not include kurtosis values in the tables. Additionally, I included qualitative content analysis to highlight specific items that had unusual response distributions or may not be suitable for operational use.

1. Epworth Sleepiness Scale

The ESS (Johns, 1991) asked participants to rate the chance of dozing under a variety of situations. Table 4 summarizes the responses received on the ESS in this study. The fourth and eighth items were scenarios in a car that sailors at sea are not likely to have encountered in the past week. The directions asked participants to estimate how they would respond even if they had not encountered the given scenarios. The responses to the sixth

and eight items were skewed to the right. Their mean values were more than one standard deviation below the midpoint of the range of 1.5.

Table 4. Epworth Sleepiness Scale Item Distribution

Item	N	Mean	SD	Skew
How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired?				
1. Sitting and reading	534	1.71	0.97	-0.18
2. Watching TV	536	1.41	0.95	0.16
3. Sitting inactive in a public place (e.g., a theater or a meeting)	536	1.28	1.04	0.27
4. As a passenger in a car for an hour without a break	535	1.51	1.12	0.00
5. Lying down to rest in the afternoon when circumstances permit	532	2.15	0.92	-0.80
6. Sitting and talking to someone	536	0.46	0.68	1.44
7. Sitting quietly after lunch without alcohol	536	1.31	0.99	0.23
8. In a car, while stopped for a few minutes in traffic	535	0.45	0.75	1.63

Notes: Item scores range from 0 to 3. Higher scores indicate greater sleepiness. Items are from Johns (1991).

2. Insomnia Severity Index

The ISI (Bastien et al., 2001), measured the severity of participants' symptoms of insomnia. Table 5 summarizes the responses received on the ISI in this study. All items exhibited a normal response distribution. The fourth item was the only one with a mean value above the midpoint response of 2. Even though other symptoms of insomnia severity are not rated as high, participants tended to be dissatisfied with their current sleep pattern.

Table 5. Insomnia Severity Index Item Distribution

Item	N	Mean	SD	Skew
1. Difficulty falling asleep	535	1.30	1.12	0.53
2. Difficulty staying asleep	535	1.50	1.14	0.27
3. Problems waking up too early	535	1.48	1.19	0.28
4. How satisfied/dissatisfied are you with your current sleep pattern?	537	2.28	0.99	-0.10
5. How noticeable to others do you think your sleep problem is in terms of impairing the quality of your life?	537	1.34	1.06	0.53
6. How worried/distressed are you about your current sleep problem?	537	1.24	1.06	0.50
7. To what extent do you consider your sleep problem to interfere with your daily functioning currently?	536	1.54	1.11	0.37

Notes: Item scores range from 0 to 4. Higher scores indicate a greater level of insomnia. Items are from Bastien, et al. (2001).

3. Pittsburgh Sleep Quality Index

The PSQI (Buysse et al., 1989) also looked at sleep-related issues. The first four questions had fill in the blank response options for entering bedtime, minutes to fall asleep, getting up time, and actual hours of sleep. The scoring algorithm assumed one sleep period per night, but some respondents reported two bedtimes and two getting up times. This may have occurred because the participant had a nighttime watch rotation and would sleep before and after standing watch. In the PSQI, higher component scores indicate a negative outcome for each component. An overall PSQI score greater than 5 on a range of 0 to 21 is an indication of a poor sleeper. In this sample, 77.1% (n=442) of participants were classified as poor sleepers. The PSQI scoring algorithm calculates seven component scores, as shown in Table 6, from the individual item responses. For this sample, the sleep duration component has the highest score and the use of sleeping medication component had the lowest score.

Table 6. Pittsburgh Sleep Quality Index Component Distribution

Component	N	Mean	SD	Skew
1. Subjective sleep quality	521	1.43	0.68	0.28
2. Sleep latency	521	1.34	1.05	0.19
3. Sleep duration	521	1.88	0.86	-0.61
4. Habitual sleep efficiency	521	0.83	1.07	0.99
5. Sleep disturbances	521	1.30	0.54	0.45
6. Use of sleeping medication	521	0.33	0.83	2.39
7. Daytime dysfunction	521	1.12	0.72	0.39

Notes: Component scores range from 0 to 3. Higher scores indicate a negative outcome for each component. Component descriptions are from Buysse, et al. (1989).

4. Connor-Davidson Resilience Scale

The CD-RISC (Campbell-Sills & Stein, 2007) used ten items to assess the participants' resilience. Table 7 summarizes the responses received on the CD-RISC in this study. All items had a similar response distribution with just a slight skew to the left. This indicates that a majority of participants believe they are exhibiting resilient behaviors.

Table 7. Connor-Davidson Resilience Scale Item Distribution

Item	N	Mean	SD	Skew
1. I am able to adapt when changes occur	537	2.85	0.80	-0.57
2. I can deal with whatever comes my way	536	2.92	0.78	-0.50
3. I try to see the humorous side of things when I am faced with problems	537	2.79	0.94	-0.58
4. Having to cope with stress can make me stronger	536	2.56	0.89	-0.34
5. I tend to bounce back after illness, injury, or other hardships	537	2.90	0.85	-0.56
6. I believe I can achieve my goals, even if there are obstacles	537	3.11	0.83	-0.91
7. Under pressure, I stay focused and think clearly	537	2.68	0.87	-0.46
8. I am not easily discouraged by failure	536	2.55	1.01	-0.41
9. I think of myself as a strong person when dealing with life's challenges and difficulties	537	2.96	0.89	-0.88
10. I am able to handle unpleasant or painful feelings like sadness, fear, and anger	537	2.79	1.01	-0.67

Notes: Item scores range from 0 to 4. Higher scores indicate greater resilience. Items are from Campbell-Sills & Stein (2007).

5. Response to Stressful Experiences Scale

The RSES (De la Rosa et al., 2016) had four items, and like the CD-RISC, it measured resilience. Table 8 summarizes the responses received on the RSES in this study. All four items had similar response distributions and had a slight left skew. The second item in the RSES and the fifth item in the CD-RISC were similar in both wording and response distributions (i.e., both items address bouncing back from stressful events). This is another indication that a majority of participants are exhibiting resilient behaviors.

Table 8. Response to Stressful Experiences Item Distribution

Item	N	Mean	SD	Skew
During and after life's most stressful events, I tend to…				
1. …find a way to do what's necessary to carry on	537	3.09	0.79	-0.60
2. …know I will bounce back	537	3.04	0.84	-0.73
3. …learn important and useful life lessons	537	2.99	0.89	-0.80
4. …practice ways to handle it better next time	537	2.84	0.95	-0.63

Notes: Item scores range from 0 to 4. Higher scores indicate greater resilience. Items are from De la Rosa et al. (2016).

6. Brief COPE

The BC (Carver, 1997) measured the use of nine coping techniques with two items for each technique. Table 9 summarizes the responses received on the BC in this study. For all nine techniques, both items had the same quartile distribution. The three coping techniques with the greatest use were Active (items 1 and 5), Reframing (items 8 and 13), and Planning (items 10 and 16). The two least used coping techniques were Denial (items 2 and 6) and Behavioral (or Giving Up) (items 4 and 12). Participants tended to use constructive coping techniques more often than destructive coping techniques.

Table 9. Brief COPE Item Distribution

Item	N	Mean	SD	Skew
1. I've been concentrating my efforts on doing something about the situation I'm in	538	2.76	0.92	-0.21
2. I've been saying to myself "this isn't real"	537	1.54	0.87	1.49
3. I've been getting emotional support from others	534	1.99	0.91	0.52
4. I've been giving up trying to deal with it	536	1.66	0.92	1.24
5. I've been taking action to try to make the situation better	537	2.85	0.87	-0.37
6. I've been refusing to believe that is has happened	534	1.38	0.71	1.79
7. I've been getting help and advice from other people	537	2.22	0.96	0.31
8. I've been trying to see it in a different light, to make it seem more positive	538	2.65	0.95	-0.27
9. I've been criticizing myself	536	2.29	1.04	0.27
10. I've been trying to come up with a strategy about what to do	538	2.73	0.94	-0.29
11. I've been getting comfort and understanding from someone	538	2.10	0.95	0.40
12. I've been giving up the attempt to cope	538	1.50	0.80	1.55
13. I've been looking for something good in what is happening	538	2.75	0.95	-0.33
14. I've been trying to find comfort in my religion or spiritual beliefs	538	1.92	1.07	0.75
15. I've been trying to get advice or help from other people about what to do	537	2.18	0.98	0.34
16. I've been thinking hard about what steps to take	538	2.60	0.99	-0.15
17. I've been blaming myself for things that happened	538	1.92	1.03	0.72
18. I've been praying or meditating	538	1.92	1.07	0.79

Notes: Item scores range from 1 to 4. Higher scores indicate greater use of coping strategy. Items are derived from Carver (1997).

7. New General Self-Efficacy Scale

The NGSES (Chen et al., 2001) measured respondents' belief in their ability to succeed. Table 10 summarizes the responses received on the NGSES. Item 1 in the NGSES and item 6 from the CD-RISC had similar wording and similar response distributions (i.e.,

both items concern the belief in the ability to achieve goals). All items had a mean value near 4.0 and a left skew. This indicated that a majority of participants were confident they would succeed. For item 7, which asks whether respondents believe they perform most tasks very well compared to other people, only 3.9% (n=21) disagreed or strongly disagreed compared with 74.5% (n=400) who agreed or strongly agreed.

Table 10. New General Self-Efficacy Scale Item Distribution

Item	N	Mean	SD	Skew
1. I will be able to achieve most of the goals I have set for myself	537	4.05	0.80	-1.00
2. When facing difficult tasks, I am certain that I will accomplish them	536	3.98	0.73	-0.64
3. In general, I think I can obtain outcomes that are important to me	536	4.13	0.68	-0.77
4. I believe I can succeed at most any endeavor to which I set my mind	537	4.17	0.76	-0.98
5. I will be able to successfully overcome many challenges	537	4.11	0.75	-0.87
6. I am confident that I can perform effectively on many different tasks	536	4.10	0.76	-0.75
7. Compared to other people, I can perform most tasks very well	537	3.97	0.83	-0.60
8. Even when things are tough, I can perform quite well	537	4.02	0.80	-0.83

Notes: Item scores range from 1 to 5. Higher scores indicate greater self-efficacy. Items are from Chen et al. (2001).

8. Brief Self-Control Scale

The BSS (Tangney et al., 2004) measured participants' level of self-control. Table 11 summarizes the responses received on the BSS in this study. All 13 items showed a normal distribution with very little skew. Nine of these items are reversed-scored so higher scores indicated greater self-control.

Table 11. Brief Self-Control Scale Item Distribution

Item	N	Mean	SD	Skew
1. I am good at resisting temptation	535	3.39	1.03	-0.34
2. I have a hard time breaking bad habits (R)	536	2.78	1.09	0.30
3. I am lazy (R)	535	3.51	1.06	-0.29
4. I say inappropriate things (R)	536	2.96	1.20	0.01
5. I do certain things that are bad for me, if they are fun (R)	536	3.13	1.12	-0.08
6. I refuse things that are bad for me	535	3.08	1.06	-0.19
7. I wish I had more self-discipline (R)	535	2.93	1.15	0.14
8. People would say that I have iron self-discipline	535	2.82	1.04	0.19
9. Pleasure and fun sometimes keep me from getting work done (R)	535	3.47	1.02	-0.34
10. I have trouble concentrating (R)	536	3.17	1.09	-0.16
11. I am able to work effectively toward long-term goals	535	3.68	0.92	-0.47
12. Sometimes I can't stop myself from doing something, even if I know it is wrong (R)	536	3.64	1.08	-0.57
13. I often act without thinking through all the alternatives (R)	536	3.46	1.11	-0.44

Notes: Item scores range from 1 to 5. Higher scores indicate greater self-control. Items marked (R) are reverse scored. Items are from Tangney et al. (2004).

9. Life Orientation Test-Revised

The LOT-R (Scheier et al., 1994) measured the optimism of the participants. Table 12 summarizes the responses received on the LOT-R in this study. All six items had normal responses, but there was a distinct difference between the positively worded items (1, 3, and 6) and the reverse-coded, negatively worded items. Respondents were more likely to agree with a positively worded item than disagree with a negatively worded item.

Table 12. Life Orientation Test-Revised Item Distribution

Item	N	Mean	SD	Skew
1. In uncertain times, I usually expect the best	539	2.36	1.07	-0.40
2. If something can go wrong for me, it will (R)	537	2.03	1.06	0.01
3. I am always optimistic about my future	538	2.63	0.96	-0.46
4. I hardly ever expect things to go my way (R)	536	2.07	1.04	0.12
5. I rarely count on good things happening to me (R)	538	2.11	1.07	-0.01
6. Overall, I expect more good things to happen to me than bad	538	2.49	1.01	-0.24

Notes: Item scores range from 0 to 4. Higher scores indicate greater optimism. Items marked (R) are reversed scored. Items are from Scheier et al. (1994).

10. Thought Control Questionnaire

The TCQ (Wells & Davies, 1994) measured how people dealt with unpleasant or unwanted thoughts. Table 13 summarizes the responses received on the TCQ. Compared to the other scales used in the SET, the directions for the TCQ were unnecessarily long and slightly misleading. This wordy introduction may have been necessary for the original TCQ, which measured five separate techniques, but this shortened form of the TCQ only had items related to the technique of positive reappraisal. All six items had similar response distributions with means near the midpoint of 2.5.

Table 13. Thought Control Questionnaire Item Distribution

Item	N	Mean	SD	Skew
When I experience an unpleasant/unwanted thought:				
1. I focus on the thought	537	2.35	0.81	0.35
2. I challenge the thought's validity	536	2.51	0.81	0.01
3. I analyze the thought rationally	537	2.74	0.80	-0.19
4. I try to reinterpret the thought	537	2.62	0.83	-0.15
5. I try a different way of thinking about it	538	2.67	0.81	-0.12
6. I question the reasons for having the thought	536	2.75	0.88	-0.17

Notes: Item scores range from 1 to 4. Higher scores indicate greater use of positive reappraisal. Items are derived from Wells & Davies (1994).

11. Cognitive Emotion Regulation Questionnaire

The CERQ (Garnefski & Kraaij, 2007) also measured the use of positive reappraisal. Table 14 summarizes the responses received on the CERQ in this study. Like the TCQ, it was part of a larger questionnaire that was shortened to focus on one technique. There was nothing unusual in the response distribution of the individual items, but the overall CERQ scores showed a multi-modal pattern with disproportionate responses at summed scores of 8, 12, 16, and 20. Further investigation revealed that over 40% of participants chose the same response for all four items.

Table 14. Cognitive Emotion Regulation Questionnaire Item Distribution

Item	N	Mean	SD	Skew
1. I can learn something from the situation	538	3.43	1.09	-0.06
2. I think I can become a stronger person as a result of what has happened	538	3.52	1.09	-0.20
3. I think that the situation also has its positive sides	538	3.33	1.11	-0.07
4. I look for the positive sides to the matter	538	3.43	1.11	-0.16

Notes: Item scores range from 1 to 5. Higher scores indicate greater use of positive reappraisal. Items are derived from Garnefski & Kraaij (2007).

12. Unit Social Support Scale

The USS (Vogt et al., 2013) used 12 items to measure the level of unit social support. Table 15 summarizes the responses received on the USS in this study. All items had a normal response distribution with mean values near the middle of the scale range. Items 3–6 focused on support from fellow unit members. There were similarly worded items for each of these that focused on support from unit leaders. For example, item 4 and item 12 both ask if the respondent feels valued by either unit members or unit leaders. There was no clear difference between the responses on items concerning leaders and fellow unit members. This indicates that, overall, participants felt that leaders in their command were providing just as much support as the other members of the command.

Table 15. Unit Social Support Scale Item Distribution

Item	N	Mean	SD	Skew
1. My unit/command is like a family to me	537	3.10	1.11	-0.27
2. People in my unit are trustworthy	537	3.23	0.96	-0.51
3. My fellow unit members appreciate my efforts	536	3.35	0.93	-0.72
4. I feel valued by my fellow unit members	537	3.30	0.98	-0.53
5. Members of my unit are interested in my well-being	537	3.36	0.94	-0.67
6. My fellow unit members are interested in what I think and how I feel about things	537	3.13	0.99	-0.43
7. My unit leader(s) are interested in what I think and how I feel about things	537	3.09	1.08	-0.31
8. I feel like my efforts really count to the leaders in my unit	537	3.17	1.07	-0.48
9. My service is appreciated by the leaders in my unit	536	3.19	1.05	-0.50
10. I could go to unit leaders for help if I had a problem or concern	537	3.36	1.12	-0.55
11. The leaders of my unit are interested in my personal welfare	537	3.27	1.09	-0.42
12. I feel valued by the leaders of my unit	537	3.21	1.11	-0.37

Notes: Item scores range from 1 to 5. Higher scores indicate greater unit support. Items are from Vogt et al. (2013).

13. NCCOSC Morale Questions

NCCOSC created two items to measure morale. Table 16 summarizes the responses received on these morale questions. Response options ranged from Very Low (0) to Very High (4). On average, respondents perceived their level of morale to be higher than the command level of morale. Nearly 50% (n=265) of participants rated their morale higher than the command and only 13% (n=72) rated their morale lower than the command.

Table 16. NCCOSC Morale Item Distribution

Item	N	Mean	SD	Skew
1. Your personal morale	535	2.18	1.04	-0.19
2. Morale in your command	535	1.64	1.02	0.03

Notes: Item scores range from 0 to 4. Higher scores indicate greater levels of morale.

14. Satisfaction with Life Scale

The SWLS (Diener et al., 1985) used five items to measure the participants' level of satisfaction. Table 17 summarizes the responses received on the SWLS in this study. Unlike most of the other SET scales, which used a four or five-point Likert scale, the SWLS used a seven-point Likert response scale. The first four items had normal response distributions, but the fifth item had a uniform distribution (i.e., each response option was endorsed at a similar frequency). Overall, 57% (n=309) of participants were satisfied with life compared to 33% (n=179) dissatisfied.

Table 17. Table N. Satisfaction with Life Scale Item Distribution

Item	N	Mean	SD	Skew
1. In most ways my life is close to my ideal	537	4.18	1.60	-0.37
2. The conditions of my life are excellent	537	4.28	1.51	-0.30
3. I am satisfied with life	537	4.72	1.61	-0.54
4. So far I have gotten the important things I want in life	536	4.75	1.61	-0.58
5. If I could live my life over, I would change almost nothing	537	3.96	1.93	-0.06

Notes: Item scores range from 1 to 7. Higher scores indicate higher satisfaction. Items are from Diener et al. (1985).

15. Short Form-36 General Health Question

The SET included only one item from the SF-36 (Ware & Sherbourne, 1992) to assess participants' overall level of health. Table 18 summarizes the responses received on this question. This scale had the most number of missing responses. The response options were listed left to right in descending order from "Excellent" to "Poor" (i.e., most favorable response listed first). This was contrary to other SET scales, which listed responses in an ascending order (i.e., most favorable response listed last).

Table 18. SF-36 General Health Questionnaire Item Distribution

Item	N	Mean	SD	Skew
1. In general, would you say your health is:	520	2.37	0.89	0.07

Notes: Item scores range from 0 to 4. Higher scores indicate better health. Item is derived from Ware & Sherbourne (1992).

16. Perceived Stress Scale

The PSS (Cohen & Williamson, 1988) used four items to measure the participant's perceived stress. Table 19 summarizes the responses received on the PSS in this study. The wording of the items was redundant as all items began with the phrase "In the past month, how often have you felt…" Examination of the item response distributions showed that items 2 and 3 were the only items in the entire SET battery for which a response option received zero endorsements. Specifically, not a single respondent selected "Very Often," which corresponds to a score of 0 for these items. Observationally, these items are the only two on the PSS that are reverse coded. It is also notable that responses to item 2 exhibited a substantive right skew, indicating respondents favored the lower end of the response scale for this item.

Table 19. Perceived Stress Scale Item Distribution

Item	N	Mean	SD	Skew
In the past month, how often have you felt …				
1. …that you were unable to control the important things in your life?	535	2.02	1.03	0.02
2. …confident about your ability to handle your personal problems? (R)	535	1.54	0.74	1.32
3. …that things were going your way? (R)	535	1.88	0.82	0.68
4. …difficulties were piling up so high that you could not overcome them?	534	1.70	1.04	0.19

Notes: Item scores range from 0 to 4. Higher scores indicate greater perceived stress. Items marked (R) are reversed scored. Items are derived from Cohen & Williamson (1988).

17. Primary Care PTSD Screen

The PC-PTSD (Prins et al., 2003) measured the number of PTSD symptoms experienced by respondents. Table 20 summarizes the responses received on the PC-PTSD in this study. The scale developers recommended a cutoff score of 3 or more for further screening of a PTSD diagnosis. In this study, 24.3% (n=130) of participants reached this cutoff; furthermore, 54.5% (n=292) of the sample reported having experienced at least one symptom of PTSD in the past month. This scale used a binary response instead of a Likert response as used by the other scales in the SET.

Table 20. Primary Care PTSD Screen Item Distribution

Item	N	Mean	SD	Skew
In your life, have you ever had any experience that was so frightening, horrible, or upsetting that, in the past month, you:				
1. Have had nightmares about it or thought about it when you did not want to?	536	0.30	0.46	0.87
2. Tried hard not to think about it or went out of your way to avoid situations that reminded you of it?	536	0.36	0.48	0.59
3. Were constantly on guard, watchful, or easily startled?	536	0.25	0.44	1.14
4. Felt numb or detached from others, activities, or your surroundings?	536	0.39	0.49	0.44

Notes: Item scores range from 0 to 1. Higher scores indicate a greater risk of PTSD. Items are from Prins et al. (2003).

18. Patient Health Questionnaire

The PHQ (Kroenke et al., 2009) had two items to measure depression and two items to measure anxiety. Table 21 summarizes the responses received on the PHQ in this study. All four items had a response distribution with a skew to the right. In this sample, 22.4% (n=120) of all respondents selected “Not at all” for all four items.

Table 21. Patient Health Questionnaire Item Distribution

Item	N	Mean	SD	Skew
Over the past two weeks, how often have you been bothered by any of the following problems?				
1. Little interest or pleasure in doing things	535	0.85	0.90	0.77
2. Feeling down, depressed or hopeless	536	0.70	0.87	1.04
3. Feeling nervous, anxious or on edge	536	0.86	0.91	0.83
4. Not being able to stop or control worrying	536	0.74	0.90	1.02

Notes: Item scores range from 0 to 3. Higher scores indicate greater levels of depression and anxiety. Items are from Kroenke et al. (2009).

C. SCALE LEVEL ANALYSIS

I assessed the internal consistency for 16 of the 18 SET scales by calculating Cronbach's alpha. NCCOSC and SF-36 had only two and one items respectively so the calculation was not applicable to these measures. As shown in Table 22, I compared the current internal consistency values with those in previously published research. In some cases, the authors conducted multiple studies so both reported values were included. Nunnally (1978) recommends a Cronbach's alpha of at least 0.70 for basic research. Only the measured value for the PSQI was below this threshold. All other scales exhibited high levels of internal consistency.

Table 22. Comparison of Cronbach's Alpha with Published Results

Scale	Measured Cronbach α	Reported Cronbach α	Source
ESS	0.82	0.88/0.73	Johns (1992)
ISI	0.82	0.90/0.91	Morin, Belleville, Bélanger, Ivers (2011)
PSQI	0.65	0.83	Buysse et al. (1988)
CD-RISC	0.90	0.85	Campbell-Sills & Stein (2007)
RSES	0.83	0.88/0.73	De La Rosa et al. (2016)
BC	0.84	N/A	(SET only used a partial scale)
NGSES	0.93	0.86/0.90	Chen et al. (2001)
BSS	0.83	0.83/0.85	Tangney et al. (2004)
LOT-R	0.84	0.78	Scheier et al. (1994)

Scale	Measured Cronbach α	Reported Cronbach α	Source
TCQ	0.80	0.67	Wells & Davies (1994)
CERQ	0.92	0.85/0.84	Garnefski & Kraaij (2007)
USS	0.95	0.96	Vogt et al. (2013)
NCCOSC	N/A	N/A	(Only 2 items in this scale)
SWLS	0.88	0.87	Diener et al. (1985)
SF-36	N/A	N/A	(Only 1 item in this scale)
PSS	0.70	0.60	Cohen & Williamson (1988)
PC-PTSD	0.80	N/A	(Not measured by previous research)
PHQ	0.84	0.85	Kroenke et al. (2009)

D. FACTOR ANALYSIS

I performed a factor analysis on 17 of the 18 scales in the SET battery. As with internal consistency, it was not applicable to perform a factor analysis on the SF-36 because it only included one item. For most of the scales, previous researchers have conducted a factor analysis. In those cases, I compared the number of factors in the SET data to the number of reported factors and, when available, the associated factor loadings. I listed items in descending order based on the factor loadings from the SET data. Hair, Tatham, Anderson, and Black (1998) give recommendations for the minimum factor loading based on the size of the sample. They recommended a minimum loading of 0.30 for sample sizes of 350; therefore, I chose 0.30 as the cutoff value for factor loadings in the current study. If a particular item did not load to at least 0.30 on any factor, I report the highest loading for that item.

1. Epworth Sleepiness Scale

Johns (1992) performed a factor analysis on the ESS scores for two diverse samples: medical students and patients suffering from sleep apnea. In both cases, Johns found that all eight items of the ESS loaded onto one factor. For the SET data, I also found only one factor. This factor had an eigenvalue of 3.53 compared to Johns reported eigenvalues of 3.95 for patients and 2.07 for students. These results confirm that one factor

of daytime sleepiness is appropriate for the ESS. Table 23 compares the factor loadings for this study with both studies by Johns (1992).

Table 23. Factor Loadings for Epworth Sleepiness Scale

Item No.	Factor Loadings:		
	SET	**Patients**	**Students**
3	0.68	0.77	0.62
4	0.66	0.68	0.54
2	0.62	0.59	0.49
1	0.61	0.73	0.55
7	0.60	0.76	0.64
8	0.57	0.73	0.37
6	0.54	0.73	0.25
5	0.51	0.53	0.49

Note: Factor loadings for patients and students from Johns (1992).

2. Insomnia Severity Index

Bastien et al. (2001) performed a factor analysis on the ISI given to insomnia patients. Their research identified three factors that they called “impact” (items 5, 6, and 7), “severity” (items 1, 2, and 3), and “satisfaction” (items 1, 4, and 6). In their analysis, items 1 and 6 loaded on two separate factors. Savard, Savard, Simard, and Ivars (2004) used the ISI with a group of cancer patients. They found a two-factor solution with items 1, 2, 3, and 4 on one factor and items 5, 6, and 7 on the other factor. The scree plot produced in the current study indicated that a three-factor solution was acceptable, but only two factors had eigenvalues greater than 1.0. In my three-factor solution, I found items 4, 6, and 7 on the first factor; items 1, 2, and 3 on the second factor; and item 5 on the third factor. Item 5 is the only item that looks at the impact of insomnia on others (i.e., all other items are related to the impact of insomnia on the individual). In my two-factor solution, items 4, 5, 6, and 7 loaded on the first factor with items 1, 2, and 3 on the second factor. There is no consensus outcome for item 4, but across studies the data support items 1, 2,

and 3 belonging to one factor and items 5, 6, 7 belonging to a separate factor. Table 24 shows the factor loadings for both the two-factor and three-factor models.

Table 24. Factor Loadings for Insomnia Severity Index

Item No.	Two-Factor Model		Three-Factor Model		
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3
7	0.82		0.75		
6	0.76		0.74		
5	0.71		0.47		0.87
4	0.66		0.66		
2		0.89		0.89	
1		0.61		0.61	
3		0.53		0.53	

3. Pittsburgh Sleep Quality Index

Previous studies investigating the factor structure of the PSQI performed the analysis on the seven component scores instead of the individual items. Both Cole et al. (2006) and Burkhalter et al. (2010) identified a three-factor model for the PSQI. They placed components 3 and 4 in the sleep efficiency factor, components 1, 2, and 6 in the sleep quality factor, and components 5 and 7 in the daily disturbance or daytime function factors. In this sample, two factors had eigenvalues greater than 1.0, but the third factor had an eigenvalue of 0.98. The scree plot suggested a two-factor model was appropriate. Both models did a poor job of explaining the variance in the data with the two-factor model accounting for 35.9% of variance and the three-factor model accounting for 39.2% of variance. Table 25 shows the factor loadings for both models created from the data in this study.

Table 25. Factor Loadings for Pittsburgh Sleep Quality Index Components

Component No.	Two-Factor Model:		Three -Factor Model:		
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3
3	0.98		0.76	0.38	
4	0.37		0.47		
5		0.56			0.59
2		0.53			0.45
1	0.40	0.48		0.56	
7		0.36		0.62	
6		0.31			0.36

4. Connor-Davidson Resilience Scale

Campbell-Sills and Stein (2007) developed the ten-item version of CD-RISC in order to have only one factor. I also found a one-factor model to be the best fitting option. Table 26 shows the factor loadings for both studies. These results confirm one factor of resilience is appropriate for this shortened version of CD-RISC.

Table 26. Factor Loadings for Connor-Davidson Resilience Scale

Item No.	Factor Loadings:	
	SET	Reported
9	0.79	0.74
6	0.75	0.63
10	0.72	0.57
2	0.72	0.72
5	0.72	0.61
4	0.70	0.58
7	0.70	0.62
1	0.69	0.44
8	0.66	0.63
3	0.49	0.46

Note: Reported factor loadings from Campbell-Sills and Stein (2007).

5. Response to Stressful Experiences Scale

De La Rosa et al. (2016) found their shortened, four-item version of the RSES to be unidimensional even though it covers four domains related to resilience. I performed a factor analysis that resulted in a one-factor solution. This single factor of resilience had an eigenvalue of 2.68 and explained 56% of the variance. Table 27 shows the factor loadings for the RSES in this study; De La Rosa et al. (2016) did not report factor loadings.

Table 27. Factor Loading for Response to Stressful Experiences Scale

Item No.	Factor 1
2	0.82
1	0.78
3	0.74
4	0.65

6. Brief COPE

Previous researchers have performed a factor analysis on the 28-item version of BC that has two items for each of 14 coping techniques. Carver (1997) identified nine factors with eigenvalues greater than 1.0 in a study on coping for hurricane survivors. Hastings et al. (2005) identified four factors using the scree test in a study of coping for parents of autistic children. The two studies disagreed on which coping techniques loaded together on the factors. In the SET, only 18 items for nine coping techniques were used for the BC. I found the best fitting solution to be a five-factor model. Table 28 shows the factor loadings for all items and the associated coping technique. These loadings were more similar to Carver's findings in that Planning, Active Coping, and Positive Reframing were loaded on the same factor. Additionally, Carver found that Instrumental Support and Emotional Support were loaded on the same factor. Hastings and colleagues found that Behavioral Disengagement and Self-Blame loaded on the same factor. I found that Behavioral Disengagement primarily loaded with Denial, but it also partially loaded on the

fifth factor with Self-Blame. Additionally, Hastings et al. found Planning and Active Coping to load on the same factor as Instrumental Support and Emotional Support.

Table 28. Factor Loadings for Brief COPE

Item No.	**Factor Loadings:**				
(Coping Technique)	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
10 (Planning)	0.72				
5 (Active)	0.69				
16 (Planning)	0.67				
13 (Positive Reframing)	0.64				
1 (Active)	0.62				
8 (Positive Reframing)	0.60				
7 (Instrumental Support)		0.78			
15 (Instrumental Support)		0.74			
11 (Emotional Support)		0.70			
3 (Emotional Support)		0.69			
6 (Denial)			0.71		
2 (Denial)			0.71		
12 (Behavioral Disengagement)			0.49		0.37
4 (Behavioral Disengagement)			0.49		0.30
18 (Religion)				0.96	
14 (Religion)				0.80	
17 (Self-blame)					0.73
9 (Self-blame)					0.70

7. New General Self-Efficacy Scale

Chen et al. (2001) created the NGSES to reduce the General Self-Efficacy Scale from three factors down to a single factor. I found the best solution to include a single factor of self-efficacy that explains 65.3% of the variance. Table 29 shows the factor loadings for this one-factor model; Chen et al. (2001) did not report factor loadings.

Table 29. Factor Loadings for New General Self-Efficacy Scale

Item No.	Factor Loadings
5	0.88
4	0.87
3	0.84
6	0.83
2	0.82
1	0.77
8	0.76
7	0.68

8. Brief Self-Control Scale

Tangney et al. (2004) developed the BSS to be a unidimensional measure of self-control. Maloney, Grawitch, and Barber (2012) concluded a two-factor model was best. However, their two-factor model used only eight of the 13 items in the BSS. I investigated several models because there was no agreement between the eigenvalues and the scree plot. The factor analysis had three factors with eigenvalues greater than 1.0 and a fourth factor that had an eigenvalue of 0.99. The scree plot had an elbow at two factors. The three-factor model had one item that loaded on all three factors, five items that loaded on two factors, and one item that did not load above 0.30 on any factor. The two-factor model only had three items that loaded on both factors and all items loaded on at least one factor. Factor 1 is associated with self-discipline and factor two is associated with impulse control. Both models did a poor job of explaining the variance in the data with the two-factor model accounting for 34.3% of variance and the three-factor model accounting for 40.7% of variance. Table 30 shows the factor loadings for this two-factor model.

Table 30. Factor Loadings for the Brief Self-Control Scale

Item No.	Factor Loadings: Factor 1	Factor 2
7	0.61	
10	0.52	
9	0.51	
3	0.50	
12	0.47	0.45
8	0.44	
13	0.44	0.37
2	0.42	0.36
1	0.37	
11	0.31	
5		0.89
4		0.65
6		0.42

9. Life Orientation Test-Revised

Scheier et al. (1994) identified a single factor that accounted for 48.1% of the variance with the LOT-R. They also performed a CFA to test a one-factor model versus a two-factor model. I created a one-factor model that had similar factor loadings and accounted for 46.8% of the variance. However, both the scree test and eigenvalue greater-than-one rule suggested a two-factor model that explained 58.6% of the variance. In the two factor model, the negatively worded (and reversed scored) items loaded on one factor and the positively worded items loaded on the second factor. This loading pattern replicates the two-factor model reported in Scheier et al. (1994). Table 31 shows a comparison of the one-factor models and the two-factor model; Scheier et al. (1994) did not report factor loadings for their two-factor model.

Table 31. Factor Loadings for Life Orientation Test-Revised

Item	One-Factor Models:		Item	Two-Factor Model:	
No.	SET	Reported	No.	Factor 1	Factor 2
5	0.80	0.74	4	0.82	
4	0.79	0.79	5	0.76	
2	0.71	0.66	2	0.66	
6	0.70	0.72	3		0.77
1	0.52	0.58	1		0.62
3	0.52	0.66	6	0.47	0.58

Note: Reported factor loadings from Scheier et al. (1994).

10. Thought Control Questionnaire

The TCQ used in the SET is a partial scale that encompasses the six items that loaded onto a single factor identified as reappraisal by Wells and Davies (1994). The SET data showed only one factor with an eigenvalue greater than 1.0. These data suggest that this shortened form of TCQ is only measuring one factor. Table 32 shows the factor loadings for both one-factor models.

Table 32. Factor Loadings for the Thought Control Questionnaire

Item No.	Factor Loadings:	
	SET	Reported
4	0.76	0.73
3	0.75	0.62
5	0.74	0.50
2	0.67	0.56
6	0.62	0.55
1	0.25	0.45

Note: Reported factor loadings from Wells and Davies (1994).

11. Cognitive Emotion Regulation Questionnaire

Garnefski and Kraaij (2007) found a nine-factor model for the CERQ. The SET only included the four items that loaded on the positive reappraisal factor. The SET data also shows all four items loading strongly on one factor that explained 73.1% of the variance. Table 33 compares the factor loadings in this study with both samples reported by Garnefski and Kraaij.

Table 33. Factor Loadings for Cognitive Emotion Regulation Questionnaire

Item No.	Factor Loadings:	
	SET	Reported
2	0.88	0.59/0.59
3	0.87	0.64/0.52
4	0.84	0.64/0.52
1	0.83	0.67/0.72

Note: Reported factor loading from Garnefski and Kraaij (2007).

12. Unit Social Support Scale

Vogt et al. (2013) did not perform a factor analysis on the USS when they updated it for the DRRI-2. In the SET data, the scree plot and eigenvalue tests both suggested a two-factor model for this scale. These two factors explained 71.3% of the variance compared with the one-factor model which explained 62.7% of the variance. Table 34 shows the factor loadings for the USS. All items loaded above 0.30 on both factors, but each item had a stronger loading on one of the factors. All six items that had a greater loading onto factor 1 correspond to the items that asked about support from unit leaders. Based on these results, I labeled the two factors measured by this scale as Support from Unit Leaders and Support from Fellow Unit Members.

Table 34. Factor Loadings for the Unit Social Support Scale

Item No.	Factor Loadings:	
	Factor 1	Factor 2
12	0.83	0.42
9	0.82	0.39
11	0.81	0.37
8	0.78	0.37
10	0.76	0.33
7	0.70	0.44
4	0.32	0.84
3	0.31	0.81
6	0.38	0.75
5	0.42	0.73
2	0.41	0.60
1	0.37	0.59

13. NCCOSC Morale Questions

The NCCOSC morale questions both loaded onto one factor of morale that explained 44.6% of the variance. Since NCCOSC developed these items specifically for the SET battery, there is no previous factor analysis with which to compare these results. Table 35 shows the factor loadings from this sample.

Table 35. Factor loadings for the NCCOSC Morale Questions

Item No.	Factor Loadings
2	0.67
1	0.67

14. Satisfaction with Life Scale

Pavot and Diener (1993) provided data from several factor analytic studies on the SWLS. In all cases, the authors reported a single factor model. The SET data also confirms that a single factor of satisfaction, which accounts for 62.2% of the variance, is applicable to the SWLS. Table 36 shows that the factor loadings are comparable with previous research.

Table 36. Factor Loadings for the Satisfaction with Life Scale

Item No.	SET	Diener et al. (1985)	Blais et al. (1989)	Arrindell et al. (1991)	Pavot et al. (1991)
3	0.86	0.83	0.74	0.85	0.82
2	0.86	0.77	0.76	0.80	0.89
1	0.81	0.84	0.84	0.84	0.83
4	0.79	0.72	0.71	0.83	0.68
5	0.60	0.61	0.68	0.76	0.78

Note: Factor loadings are as reported in Pavot & Diener (1993)

15. Short Form–36 General Health Question

I did not perform a factor analysis on this scale because it only contained one item.

16. Perceived Stress Scale

Cohen and Williamson (1988) reported a single factor for the PSS that accounted for 45.6% of the variance. The factor analysis on the SET data only had one factor with an eigenvalue greater than 1.0. This data suggest a model with a single factor related to perception of stress in the previous month. This model accounted for 38.6% of the variance. Table 37 lists the factor loadings for the PSS in this study; Cohen and Williamson (1988) did not report factor loadings.

Table 37. Factor Loadings for Perceived Stress Scale

Item No.	Factor Loadings
3	0.71
1	0.60
4	0.59
2	0.58

17. Primary Care PTSD Screen

Prins et al. (2003) acknowledged the existence of four underlying factors to PTSD (i.e., re-experiencing, numbing, avoidance, and hyperarousal), which is why they chose four items for this scale. However, they did not perform a factor analysis on the PC-PTSD scale. In the SET data, this scale only had one factor with an eigenvalue greater than 1.0. This one factor explained 51.3% of the variance. Table 38 lists the factor loadings for this single factor model.

Table 38. Factor Loadings for Primary Care PTSD Screen

Item No.	Factor Loadings
2	0.81
1	0.79
3	0.64
4	0.61

18. Patient Health Questionnaire

Kroenke et al. (2009) found a two-factor solution for the four-item PHQ. Factor 1 includes the two items related to anxiety, and factor 2 includes the two items related to depression. The factor analysis on the SET data only yielded one factor with an eigenvalue greater than 1.0, but the scree plot indicated two factors might be appropriate. The single factor model explained 57.9% of the variance, but the two-factor model explained 69.7%. Table 39 shows the two-factor model has comparable loadings to those reported by Kroenke and colleagues.

Table 39. Factor Loadings for the Patient Health Questionnaire

Item No.	SET		Reported	
	Factor 1	**Factor 2**	**Factor 1**	**Factor 2**
3	0.83		0.87	
4	0.72	0.37	0.86	0.33
1		0.81		0.90
2	0.53	0.62	0.40	0.82

Note: Reported factor loadings from Kroenke et al. (2009).

E. BATTERY-WIDE ANALYSIS OF SET

In order to improve the SET, I examined the scales in light of their item content, psychometric properties, as well as the inter-correlations between scale scores. In this section, I describe how I chose to retain or eliminate particular scales. The 18 SET scales created 153 inter-correlation values, as reported in Table 40. Overall, the direction of the correlations was consistent with expectations. That is, scales for which high scores carried desirable connotations tended to be positively correlated to one another and negatively correlated to scales that carried undesirable connotations, and vice versa. The one exception was the TCQ (high scores desirable), which had positive correlations with five of the six scales that measured negative outcomes as well as positive correlations with all of the scales that measured positive outcomes.

The magnitude of correlations ranged from 0.01 to 0.69, and the mean correlation was 0.30. For the purpose of this analysis, I defined a correlation as relatively strong if it was in the top quartile among all 153 correlations. This upper quartile included the correlations greater than or equal to 0.41. This value was not an absolute cutoff value to determine convergent validity but rather a recommendation for further examination to determine if the scale should be included in the SET battery. The red circles on Table 40 illustrate the pattern of certain scales (e.g. CD-RISC, LOT-R, CERQ, and PSS) that had relatively strong correlations with multiple other scales in the SET battery.

Table 40. Correlation of Individual Scale Scores

	ESS*	ISI*	PSQI*	CD-RISC	RSES	BC	NGSES	BSS	LOT-R	TCQ	CERQ	USS	NCCOSC	SWLS	SF-36	PSS*	PC-PTSD*	PHQ*
ESS*																		
ISI*	.34																	
PSQI*	.18	.67																
CD-RISC	-.06	-.20	-.12															
RSES	-.02	-.10	-.11	.69														
BC	.02	-.14	-.13	.41	.46													
NGSES	-.04	-.12	-.06	.63	.60	.44												
BSS	-.17	-.17	-.10	.31	.35	.29	.38											
LOT-R	-.07	-.22	-.22	.43	.43	.43	.51	.40										
TCQ	.12	.05	.01	.22	.31	.22	.23	.08	.12									
CERQ	-.02	-.12	-.14	.47	.59	.52	.50	.30	.49	.43								
USS	-.15	-.27	-.22	.33	.28	.33	.30	.22	.41	.06	.25							
NCCOSC	-.22	-.37	-.34	.37	.35	.35	.32	.27	.44	.10	.29	.55						
SWLS	-.12	-.28	-.20	.38	.36	.35	.35	.33	.48	.04	.33	.47	.53					
SF-36	-.11	-.28	-.24	.39	.35	.29	.37	.28	.39	.10	.29	.33	.38	.33				
PSS*	.15	.35	.30	-.50	-.42	-.37	-.42	-.33	-.59	-.03	-.42	-.48	-.46	-.54	-.41			
PC-PTSD*	.17	.35	.28	-.20	-.14	-.17	-.16	-.21	-.29	.06	-.13	-.21	-.18	-.25	-.25	.39		
PHQ*	.17	.42	.33	-.35	-.27	-.27	-.28	-.30	-.47	.07	-.25	-.40	-.46	-.46	-.36	.60	.48	

Note: Scales marked with a * are scales where a high score indicates a negative outcome such as sleepiness, stress, PTSD, or anxiety.

V. CONCLUSIONS AND RECOMMENDATIONS

The Navy recognizes that resilience is important to maintaining the readiness of the fleet. NCCOSC developed the SET to assess resilience in sailors and marines. This thesis sought to measure the psychometric properties of the SET and to use this information to make potential improvements. Based on the results of these analyses, I can draw several conclusion about the reliability and validity of the SET and make recommendation to reduce the length of the SET battery. I would also like to discuss limitations associated with this study and make recommendations for future research on this topic.

A. RECOMMENDATIONS TO SHORTEN THE SET BATTERY

The ESS, ISI, and PSQI are all measures that concern sleep. I first eliminated the PSQI because it had low internal consistency, and the factor analysis did not show a clear factor structure. Furthermore, the PSQI had fill-in-the-blank items for usual sleep and wake times. These items in particular were very subjective because sailors often sleep at different times every day or have multiple sleep periods. The PSQI did have a correlation of 0.67 with the ISI so it may be redundant to keep both measures. The ESS and ISI were comparable with internal consistency, but the ISI had stronger correlations with the other scales. While the items in the ISI are in general terms, the ESS items describe specific activities that may not be applicable to sailors at sea. For these reasons, I eliminated the ESS and recommend that only the ISI be included.

Both the CD-RISC and RSES are measures that concern resilience in general. They had the highest correlation between any two scales in the SET battery at 0.69. Both scales had good internal consistency, but CD-RISC had a larger Cronbach's alpha than RSES. In most cases, CD-RISC and RSES had relatively similar magnitudes of correlations to the other scales in the SET. Even though both scales are psychometrically sound, it is redundant to include both the CD-RISC and RSES. I recommend keeping CD-RISC and eliminating RSES.

The TCQ and CERQ both measure the same factor of positive reappraisal. The CERQ showed a higher internal consistency and, in all cases, had a stronger correlation

with the other SET scales. Finally, the wording of items and response options for the CERQ were clearer than the wording and response options in the TCQ. For these reasons, I recommend eliminating the TCQ.

The USS and the NCCOSC morale questions are the only two scales written specifically for a military environment, and they had a correlation of 0.55. The USS measured the support received from unit leaders and fellow unit members. The NCCOSC questions about morale measured individual morale and perception of the overall morale in the command. The individual item analysis showed there was a bias towards underestimating the morale in the command. The USS had excellent internal consistency with a Cronbach's alpha of 0.95. The USS was also developed as part of a larger Deployment Risk and Resiliency Inventory, which was specifically developed to assess resilience of military personnel. For these reasons, I recommend keeping the USS and eliminating the NCCOSC morale questions.

For the remaining scales, I looked at the correlation patterns to determine which scales had a relatively strong correlation with multiple other scales. The BC, NGSES, LOT-R, PSS, and PHQ all fell into this category, and I recommend keeping them in the SET. Similarly, I looked to eliminate scales that only had relatively strong correlations with a few scales. The BSS, SWLS, SF-36 health question, and PC-PTSD fell into this category. These scales also tended to have other weaknesses that contributed to the recommendation for exclusion. For example, the BSS had an inconsistent factor structure. The PC-PTSD had wording that could be confusing. The SWLS, SF-36 health question, and PC-PTSD had response options that were different from other scales in the SET battery. The SWLS had a seven-point Likert response, the SF-36 health question listed responses in descending order from most favorable response to least favorable response, and the PC-PTSD had a binary response. Table 41 is a summary of which scales I recommended to retain or eliminate. Based on these recommendations, the SET would be reduced from 134 items to 73 items.

Table 41. Recommended Scales to Retain or Eliminate from the SET

Scales to Retain	Scales to Eliminate
ISI	ESS
CD-RISC	PSQI
BC	RSES
NGSES	BSS
LOT-R	TCQ
CERQ	NCCOSC
USS	SWLS
PSS	SF-36
PHQ	PC-PTSD

B. CONCLUSIONS

Good analysis begins with good data. Since NCCOSC designed the SET for use with sailors, the sample data should be representative of the overall Navy population. The demographic results confirm that the sample was representative in terms of age, gender, and rank. The only discrepancies were at the junior enlisted and senior officer ranks. Both of these slight discrepancies are likely due to the sample coming from seagoing commands. Many junior enlisted, especially E1 and E2 sailors, are still at their initial training commands. Senior officers are also more likely to be on staff or shore billets instead of sea billets. Similarly, older sailors generally have more seniority and are less likely to be at sea billets. Overall, the data from the current study were generally representative of the larger Navy.

The first goal of this study was to measure the psychometric properties of the scales in the SET battery. While some scales were more reliable than others were, the results showed that all scales were generally psychometrically sound. The individual items tended to have normal response distributions. The internal consistency values were similar to those reported by previous researchers. The factor structure of most scales matched those reported by previous researchers. I attribute this result to NCCOSC choosing to use already established scales instead of creating entirely new scales with novel items. Even though

some scales were slightly modified in comparison to their original design, they performed well in this study with this population.

The second goal of this study was to recommend improvements to the SET battery because, in practice, it took participants a long time to fill out the entire questionnaire. The result of the analysis was a recommendation to shorten the SET battery from 18 scales down to nine. This reduced the number of items by over 45%. Among the remaining scales, all but the BC and PHQ have a five-point Likert response. For consistency, I would change the BC and PHQ from a four-point scale to a five-point scale. For example, the PHQ could add "About half the days" as a response option between "Several days" and "More than half the days." I would also review the instructions for each scale to make sure the time span they reference is consistent. Some ask for responses based on the past week or past month and others do not reference a specific time.

While the purpose of the SET is to assess resilience, the individual scales addressed a variety of constructs related to resilience. Among the nine scales recommended for retention, the ISI, PSS, and PHQ have a different relation to resilience than the others. Insomnia, stress, anxiety, and depression are outcomes of not being resilient as opposed to the other constructs such as coping, self-efficacy, optimism, positive reappraisal, and unit support, which are constructs that build resilience. This distinction separates the scales into two groups. This grouping aligns with the intentions of NCCOSC to identify a core group of resilience measures along with other clinical symptoms and health behaviors related to resilience. I agree that both groups of measures are important and should remain part of the SET battery.

C. LIMITATIONS

One limitation of this study is that it is only a one-time snapshot of the resilience measured by the SET. The larger research project entailed two data collection phases over the course of a single deployment period. The SET data obtained for this thesis were collected in the first phase during the first month of deployment. As such, there is no opportunity to measure the reliability of SET by retesting the same individuals at two points.

Another limitation is that the data in this study are only from sailors on surface vessels while at sea. This may not be a concern because the target audience of the SET is sailors and marines. Researchers have developed and tested many of these scales on diverse populations including university students and clinical samples. The comparison of the results of this sample with previous research is not an apples-to-apples comparison because of the difference between the sample populations. Furthermore, additional research may need to be conducted to determine if the results are applicable to populations of Navy personnel from other platforms, communities, or work environments, such as aviation, submarines, special warfare, or shore duty.

D. RECOMMENDATIONS FOR FUTURE WORK

This study lays the groundwork for many potential avenues of research on the SET and resilience. This study focused on the psychometric properties of the SET battery, but the data gathered in this sample could be used for additional research. Responses from the SET could be analyzed with data gathered as part of the larger study on human performance to compare the resilience of crewmembers on circadian versus non-circadian watch rotations. Researchers could also analyze the data with regard to other breakout groups such as age, rank, gender, ship, or department. They could also perform a closer examination of specific items from the SET battery. For example, item 4 on the ISI measured satisfaction with current sleep patterns. Instead of just examining the mean response, it might be useful to explore the watch rotations of those who answered “Very Dissatisfied” compared to those who answered “Very Satisfied.”

Furthermore, researchers could use the shortened version of the SET battery in future data collections. They could collect data from sailors in other environments such as aviation, submarines, amphibious vessels, or shore duty. A longitudinal study could examine test-retest reliability and determine how resilience changes over time. This research could be used to establish norms for this active duty population. Future studies could compare scores on the SET to relevant performance outcomes as way to establish criterion validity for the SET battery. All of these studies would further validate the usefulness of SET in assessing the resilience of Sailors and Marines.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX. SET BATTERY

Naval Postgraduate School

Date: _____ Participant ID: _____

SET Only Study Questionnaire

Instructions: *Please answer ALL questions as accurately as possible. ALL information is confidential and will be used only for research purposes.*

1. What is your age: _____ years
2. Gender (Check one ☑) ☐ Male ☐ Female
3. What is your rate: (for example, FC, HT, OS, IT, GSE) _____
4. What is your rank: (for example, E4, O2) _____
5. Department: (for example, Engineering, Operations, etc.) _____
6. Years on active duty: _____
7. How many times have you deployed: _____ Total number of months deployed: _____
8. In your previous experiences, what watch schedules did you stand (e.g., 4/8, 12/12, 5/10, 5/15, etc.)? _____
9. Which of the following nicotine products do you use? (Check ALL that apply ☑) and indicate how often)
   - ☐ Cigarettes (If YES, how often? _____)
   - ☐ Chewing tobacco/snuff (If YES, how often? _____)
   - ☐ Nicorette gum or patches (If YES, how often? _____)
   - ☐ Electronic smoke (If YES, how often? _____)
   - ☐ Other (specify): _____ (How often? _____)
10. What things affect your sleep? (Check ALL that apply ☑)
    - ☐ Not enough time to sleep
    - ☐ Noise: _____Other people _____ Noise inside berthing area _____ Noise outside berthing _____ 1MC
    - ☐ Temperature: _____Too cold ____Too hot
    - ☐ Light
    - ☐ Motion
    - ☐ Bedding Conditions: ____Bed size ____Mattress ____Pillow _____ Curtain
    - ☐ Odors
    - ☐ Other things that affect your sleep: _____
11. How many of the following caffeinated beverages do you drink **on average** each day? (Check ALL that apply ☑) and indicate daily amount)
    - ☐ Tea Servings/Cups per day: _____
    - ☐ Coffee Servings/Cups per day: _____
    - ☐ Soda/pop/soft drinks Servings/Cups per day: _____
    - ☐ Energy drinks (Monster/RedBull, etc.) Servings/Cups per day: _____
    - ☐ Other (specify): _____ How often: _____ (Example: 4 times per day)

12. Do you take any prescribed or over-the-counter medications? (Check one ☑) ☐ Yes ☐ No
If **YES**, please list all medications you take:
_____
_____

13. Do you have an exercise routine? (Check one ☑) ☐ Yes ☐ No
If **YES,** frequency: ____Daily ____Times per week (for example, 3 Times per week)
What kind of exercise routine do you do? (for example, cardio, weight lifting)
_____
How long does this routine take? (for example, 45 minutes) _____

ESS instructions: *How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired? This refers to your usual way of life in the* ***last week****. Even if you have not done some of these things recently try to work out how they would have affected you. Check ☑ the most appropriate number for each situation.*

	CHANCE OF DOZING			
	None (0)	Slight (1)	Moderate (2)	High (3)
Sitting and reading	○	○	○	○
Watching TV	○	○	○	○
Sitting inactive in a public place (e.g., a theater or a meeting)	○	○	○	○
As a passenger in a car for an hour without a break	○	○	○	○
Lying down to rest in the afternoon when circumstances permit	○	○	○	○
Sitting and talking to someone	○	○	○	○
Sitting quietly after a lunch without alcohol	○	○	○	○
In a car, while stopped for a few minutes in traffic	○	○	○	○

ISI instructions: *Please rate the severity of your insomnia symptoms during the* ***last week****. Check ☑ the most appropriate for each situation.*

	None (0)	Mild (1)	Moderate (2)	Severe (3)	Very Severe (4)
Difficulty falling asleep	○	○	○	○	○
Difficulty staying asleep	○	○	○	○	○
Problems waking up too early	○	○	○	○	○
How SATISFIED/DISSATISFIED are you with your CURRENT sleep pattern?	Very Satisfied ○	Satisfied ○	Moderately Satisfied ○	Dissatisfied ○	Very Dissatisfied ○
How NOTICEABLE to others do you think your sleep problem is in terms of impairing the quality of your life?	Not at all Noticeable ○	A Little ○	Somewhat ○	Much ○	Very Much Noticeable ○
How WORRIED/DISTRESSED are you about your current sleep problem?	Not at all Worried ○	A Little ○	Somewhat ○	Much ○	Very Much Worried ○
To what extent do you consider your sleep problem to INTERFERE with your daily functioning CURRENTLY? (i.e., daytime fatigue, mood, ability to function at work, concentration, memory, mood, etc.)	Not at all Interfering ○	A Little ○	Somewhat ○	Much ○	Very Much Interfering ○

PSQI instructions: *The following questions relate to your usual sleep habits during the* ***last week only****. Your answers should indicate the most accurate reply for the majority of days/nights in the* ***last week****. Please answer all questions.*

1. In the last week, what time have you usually gone to bed at night?	Bed Time: _____
2. During the last week, how long (in minutes) has it usually taken you to fall asleep each night	Number of Minutes:_____
3. In the last week, what time have you usually gotten up in the morning?	Getting up time:_____
4. During the last week, how many hours of actual sleep did you get at night? (this may be different than the number of hours you spent in bed.)	Hours of Sleep per Night:__

Instructions: *For each of the questions, check the one best response.*

5. During the last week, how often have you had trouble sleeping because you…	Not during the past month	Less than once a week	Once or twice a week	3 or more times a week
a) Cannot get to sleep within 30 mins	○	○	○	○
b) Wake up in the middle of the night or early morning	○	○	○	○
c) Have to get up to use the bathroom	○	○	○	○
d) Cannot breathe comfortably	○	○	○	○
e) Cough or snore loudly	○	○	○	○
f) Feel too cold	○	○	○	○
g) Feel too hot	○	○	○	○
h) Had bad dreams	○	○	○	○
i) Have pain	○	○	○	○
j) Other reason(s), please describe: _____ How often during the last week have you had trouble sleeping because of this other reason?	○	○	○	○
6. During the last week, how would you rate your sleep quality overall?	Very Good ○	Fairly Good ○	Fairly Bad ○	Very Bad ○
7. During the last week, how often have you taken medicine to help you sleep (prescribed or "over the counter"?	Not during the past week ○	Less than once a week ○	Once or twice a week ○	Three or more times a week ○
8. During the last week, how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?	○	○	○	○
9. During the last week, how much of a problem has it been for you to keep up enough enthusiasm to get things done?	Not a problem at all ○	Only a very slight problem ○	Somewhat of a problem ○	A very big problem ○

(CD-RISK) Please indicate how much you agree with the following statements as they apply to you over the **last month**. If a particular situation has not occurred recently, answer according to how you think you would have felt.					
	Not true at all	Rarely true	Sometimes true	Often true	True nearly all the time
I am able to adapt when changes occur	○	○	○	○	○
I can deal with whatever comes my way	○	○	○	○	○
I try to see the humorous side of things when I am faced with problems	○	○	○	○	○
Having to cope with stress can make me stronger	○	○	○	○	○
I tend to bounce back after illness, injury, or other hardships	○	○	○	○	○
I believe I can achieve my goals, even if there are obstacles	○	○	○	○	○
Under pressure, I stay focused and think clearly	○	○	○	○	○
I am not easily discouraged by failure	○	○	○	○	○
I think of myself as a strong person when dealing with life's challenges and difficulties.	○	○	○	○	○
I am able to handle unpleasant or painful feelings like sadness, fear, and anger	○	○	○	○	○

(RSES-4) The following statements describe how some individuals may think, feel, or act during and after the most stressful events in life. Please indicate how well each of these statements describes you during and after life's most stressful events.					
During and after life's most stressful events, I tend to….	Not at all like me	Unlike me	Sometimes like me	Like me	Very much like me
…find a way to do what's necessary to carry on	○	○	○	○	○
…know I will bounce back	○	○	○	○	○
…learn important and useful life lessons	○	○	○	○	○
…practice ways to handle it better next time	○	○	○	○	○

(BC-18) These items deal with ways that you cope with the stress in your life. Obviously, different people deal with stress in different ways, and we are interested in how YOU try to deal with it. Each item says something about a particular way of coping. Please answer to what extent you've been doing what the item says. Don't answer on the basis of whether it seems to be working or not -- just whether or not you're doing it. Using these response choices, try to rate each item separately in your mind from the others. Make your answers as true FOR YOU as you can.

	I haven't been doing this at all	I've been doing this a little bit	I've been doing this a medium amount	I've been doing this a lot
I've been concentrating my efforts on doing something about the situation I'm in	○	○	○	○
I've been saying to myself "this isn't real"	○	○	○	○
I've been getting emotional support from others	○	○	○	○
I've been giving up trying to deal with it	○	○	○	○
I've been taking action to try to make the situation better	○	○	○	○
I've been refusing to believe that it has happened	○	○	○	○
I've been getting help and advice from other people	○	○	○	○
I've been trying to see it in a different light, to make it seem more positive	○	○	○	○
I've been criticizing myself	○	○	○	○
I've been trying to come up with a strategy about what to do	○	○	○	○
I've been getting comfort and understanding from someone	○	○	○	○
I've been giving up the attempt to cope	○	○	○	○
I've been looking for something good in what is happening	○	○	○	○
I've been trying to find comfort in my religion or spiritual beliefs	○	○	○	○
I've been trying to get advice or help from other people about what to do	○	○	○	○
I've been thinking hard about what steps to take	○	○	○	○
I've been blaming myself for things that happened	○	○	○	○
I've been praying or meditating	○	○	○	○

(NGSES) Read each of the statements below and mark the response that best fits your personal belief. To what extent does each statement describe you? Indicate your level of agreement by marking the appropriate response on the right.					
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I will be able to achieve most of the goals that I have set for myself	○	○	○	○	○
When facing difficult tasks, I am certain that I will accomplish them	○	○	○	○	○
In general, I think that I can obtain outcomes that are important to me	○	○	○	○	○
I believe I can succeed at most any endeavor to which I set my mind	○	○	○	○	○
I will be able to successfully overcome many challenges	○	○	○	○	○
I am confident that I can perform effectively on many different tasks	○	○	○	○	○
Compared to other people, I can do most tasks very well	○	○	○	○	○
Even when things are tough, I can perform quite well	○	○	○	○	○

(BSS) Using the scale provided, please indicate how much each of the following statements reflects how you typically are.					
	Not at all like me	Unlike me	Sometimes like me	Like me	Very much like me
I am good at resisting temptation	○	○	○	○	○
I have a hard time breaking bad habits	○	○	○	○	○
I am lazy	○	○	○	○	○
I say inappropriate things	○	○	○	○	○
I do certain things that are bad for me, if they are fun	○	○	○	○	○
I refuse things that are bad for me	○	○	○	○	○
I wish I had more self-discipline	○	○	○	○	○
People would say that I have iron self-discipline	○	○	○	○	○
Pleasure and fun sometimes keep me from getting work done	○	○	○	○	○
I have trouble concentrating	○	○	○	○	○
I am able to work effectively toward long-term goals	○	○	○	○	○
Sometimes I can’t stop myself from doing something, even if I know it is wrong	○	○	○	○	○
I often act without thinking through all the alternatives	○	○	○	○	○

(LOT-R-6) Please be as honest and accurate as you can throughout. Try not to let your response to one statement influence your responses to other statements. There are no “correct” or “incorrect” answers. Answer according to your own feelings, rather than how you think “most people” would answer.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
In uncertain times, I usually expect the best	○	○	○	○	○
If something can go wrong for me, it will	○	○	○	○	○
I am always optimistic about my future	○	○	○	○	○
I hardly ever expect things to go my way	○	○	○	○	○
I rarely count on good things happening to me	○	○	○	○	○
Overall, I expect more good things to happen to me than bad	○	○	○	○	○

(TCQ-6) Most people experience unpleasant and/or unwanted thoughts (in verbal and/or picture form) which can be difficult to control. We are interested in the techniques that you generally use to control such thoughts. Below are a number of things that people do to control these thoughts. Please read each statement carefully, and indicate how often you use each technique. There are no right or wrong answers. Do not spend too much time thinking about each one.

When I experience an unpleasant / unwanted thought:	Never	Sometimes	Often	Almost always
I focus on the thought	○	○	○	○
I challenge the thought’s validity	○	○	○	○
I analyze the thought rationally	○	○	○	○
I try to reinterpret the thought	○	○	○	○
I try a different way of thinking about it	○	○	○	○
I question the reasons for having the thought	○	○	○	○

(CERQ) Everyone gets confronted with negative or unpleasant events now and then and everyone responds to them in his or her own way. By the following questions you are asked to indicate what you generally think, when you experience negative or unpleasant events.

	Almost never	Sometimes	Regularly	Often	Almost always
I think I can learn something from the situation	○	○	○	○	○
I think that I can become a stronger person as a result of what has happened	○	○	○	○	○
I think that the situation also has its positive sides	○	○	○	○	○
I look for the positive sides to the matter	○	○	○	○	○

(USS) The statements below are about your relationships with other military personnel. As used in these statements, the term “unit” refers to those you work with on a daily basis. Please mark how much you agree or disagree with each statement.					
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
My unit/command is like a family to me	○	○	○	○	○
People in my unit are trustworthy	○	○	○	○	○
My fellow unit members appreciate my efforts	○	○	○	○	○
I feel valued by my fellow unit members	○	○	○	○	○
Members of my unit are interested in my well-being	○	○	○	○	○
My fellow unit members are interested in what I think and how I feel about things	○	○	○	○	○
My unit leader(s) are interested in what I think and how I feel about things	○	○	○	○	○
I feel like my efforts really count to the leaders in my unit	○	○	○	○	○
My service is appreciated by the leaders in my unit	○	○	○	○	○
I could go to unit leaders for help if I had a problem or concern	○	○	○	○	○
The leaders of my unit are interested in my personal welfare	○	○	○	○	○
I feel valued by the leaders of my unit	○	○	○	○	○

(NCCOSC) Please rate each level of morale.					
	Very low	Low	Medium	High	Very high
Your personal morale	○	○	○	○	○
Morale in your command	○	○	○	○	○

(SWLS) Below are five statements with which you may agree or disagree. Indicate your agreement with each item by placing the appropriate mark after the item. Please be open and honest in your responding.							
	Strongly disagree	Disagree moderately	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
In most ways my life is close to my ideal	○	○	○	○	○	○	○
The conditions of my life are excellent	○	○	○	○	○	○	○
I am satisfied with life	○	○	○	○	○	○	○
So far I have gotten the important things I want in life	○	○	○	○	○	○	○
If I could live my life over, I would change almost nothing	○	○	○	○	○	○	○

(SF-36) In general, would you say your health is:				
Excellent	Very good	Good	Fair	Poor
○	○	○	○	○

(PSS-4) The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate how often you felt or thought a certain way.					
	Never	Almost never	Some times	Fairly often	Very often
In the past month, how often have you felt that you were unable to control the important things in your life?	○	○	○	○	○
In the past month, how often have you felt confident about your ability to handle your personal problems?	○	○	○	○	○
In the past month, how often have you felt that things were going your way?	○	○	○	○	○
In the past month, how often have you felt difficulties were piling up so high that you could not overcome them?	○	○	○	○	○

(PC-PTSD) In your life, have you ever had any experience that was so frightening, horrible, or upsetting that, in the past month, you:		
	Yes	No
Have had nightmares about it or thought about it when you did not want to?	○	○
Tried hard not to think about it or went out of your way to avoid situations that reminded you of it?	○	○
Were constantly on guard, watchful, or easily startled?	○	○
Felt numb or detached from others, activities, or your surroundings?	○	○

(PHQ-4) Over the past two weeks, how often have you been bothered by any of the following problems?				
	Not at all	Several days	More than half the days	Nearly every day
Little interest or pleasure in doing things	○	○	○	○
Feeling down, depressed or hopeless	○	○	○	○
Feeling nervous, anxious, or on edge	○	○	○	○
Not being able to stop or control worrying	○	○	○	○

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Agaibi, C. E., & Wilson, J. P. (2005). Trauma, PTSD, and resilience: A review of the literature. *Trauma, Violence, and Abuse, 6*(3), 195–216. doi:10.1177/1524838005277438

Arrindell, W. A., Meeuwesen, L., & Huyse, F. J. (1991). The Satisfaction with Life Scale (SWLS): Psychometric properties in a non-psychiatric medical outpatients sample. *Personality and Individual Differences, 12*(2), 117–123. doi:10.1016/0191-8869(91)90094-R

Babson, K. A., Blonigen, D. M., Boden, M. T., Drescher, K. D., & Bonn-Miller, M. O. (2012). Sleep quality among U.S. military veterans with PTSD: A factor analysis and structural model of symptoms. *Journal of Traumatic Stress, 25*, 665–674. doi:10.1002/jts.21757

Bastien, C. H., Vallières, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Medicine, 2*, 297–307.

Blais, M. R., Vallerand, R. J., Pelletier, L. G., & Briere, N. M. (1989). L'Echelle de satisfaction de vie: Validation Canadienne-française du "Satisfaction with Life Scale" [French-Canadian validation of the Satisfaction with Life Scale]. *Canadian Journal of Behavioral Science, 21*(2), 210–223.

Bonanno, G. A. (2004). Loss, trauma, and human resilience: have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist, 59*(1), 20–28. doi:10.1037/0003-066X.59.1.20

Bryan, C. J. (2011). The clinical utility of a brief measure of perceived burdensomeness and thwarted belongingness for the detection of suicidal military personnel. *Journal of Clinical Psychology, 67*(10), 981–992. doi:10.1002/jclp.20726

Burkhalter, H., Sereika, S. M., Engberg, S., Wirz-Justice, A., Steiger, J., & De Geest, S. (2010). Structure validity of the Pittsburgh Sleep Quality Index in renal transplant recipients: A confirmatory factor analysis. *Sleep and Biological Rhythms, 8*, 274–281. doi:10.1111/j.1479-8425.2010.00473.x

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatric Research, 28*, 193–213.

Campbell-Sills, L., & Stein, M. B. (2007). Psychometric analysis and refinement of the Connor-Davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. *Journal of Traumatic Stress, 20*(6), 1019–1028. doi:10.1002/jts.20271

Carlson, K. D., & Herdman, A. O. (2012). Understanding the impact of convergent validity on research results. *Organizational Research Methods, 15*(1), 17–32. doi:10.1177/1094428110392383

Carver, C. S. (1997). You want to measure coping but your protocol's too long: consider the Brief COPE. *International Journal of Behavioral Medicine, 4*(1), 92–100.

Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: a theoretically based approach. *Journal of Personality and Social Psychology, 56*(2), 267–283.

Chairman of the Joints Chiefs of Staff. (2011). *Chairman's Total Force Fitness Framework (CJCSI 3405.01).* Washington, DC: Joint Chiefs of Staff.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a New General Self-Efficacy Scale. *Organizational Research Methods, 4*(1), 62–83.

Cohen, S., & Williamson, G. M. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan, & S. Oskamp (Eds.), *The social psychology of health* (pp. 31–67). Newbury Park, CA: Sage Publications.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*(4), 385–396.

Cole, J. C., Motivala, S. J., Buysse, D. J., Oxman, M. N., Levin, M. J., & Irwin, M. R. (2006). Validation of a 3-factor scoring model for the Pittsburgh Sleep Quality Index in older adults. *Sleep, 29*(1), 112–116.

Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety, 18*, 76–82. doi:10.1002/da.10113

De La Rosa, G. M., Webb-Murphy, J. A., & Johnston, S. L. (2016). Development and validation of a brief measure of psychological resilience: An adaptation of the Response to Stressful Experiences Scale. *Military Medicine, 181*(3), 202–208.

Department of the Air Force. (2014). *Comprehensive Airman fitness (Air Force Instruction 90–506).* Secretary of the Air Force. Retrieved from http://static.e-publishing.af.mil/production/1/saf_mr/publication/afi90-506/afi90-506.pdf

Department of the Army. (2014). *Comprehensive soldier and family fitness (Army Regulation 350–53).* Washington, DC: Department of the Army Headquarters. Retrieved from https://www.army.mil/e2/downloads/rv7/r2/policydocs/r350_53.pdf

DeVellis, R. F. (2003). *Scale development: theory and applications* (Vols. 26, Applied Social Research Methods). Thousand Oaks, CA: SAGE.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*(1), 71–75.

Fletcher, D., & Sarkar, M. (2013). Psychological resilience: a review and critique of definitions, concepts, and theory. *European Psychologist, 18*, 12–23. doi:10.1027/1016-9040/a000124

Garnefski, N., & Kraaij, V. (2007). The Cognitive Emotion Regulation Questionnaire: psychometric features and prospective relationships with depression and anxiety in adults. *European Journal of Psychological Assessment, 23*(3), 141–149. doi:10.1027/1015-5759.23.3.141

Hair, J. F., Tatham, R. L., Anderson, R. E., & Black, W. (1998). *Multivariate data analysis* (5th ed.). London: Prentice-Hall.

Hastings, R. P., Kovshoff, H., Brown, T., Ward, N. J., Espinosa, F. D., & Remington, B. (2005). Coping strategies in mothers and fathers of preschool and school-age children with autism. *Autism, 9*(4), 377–391. doi:10.1177/1362361305056078

Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep, 14*(6), 540–545.

Johns, M. W. (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep, 15*(4), 376–381.

Johnson, D. C., Polusny, M. A., Erbes, C. R., King, D., King, L., Litz, B. T., . . . Southwick, S. M. (2011). Development and initial validation of the response to Stressful Experiences Scale. *Military Medicine, 176*(2), 161–169.

Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52–54. doi:10.5395/rde.2013.38.1.52

King, L. A., King, D. W., Vogt, D. S., Knight, J., & Samper, R. E. (2006). Deployment Risk and Resilience Inventory: A collection of measures for studying deployment-related experiences of military personnel and veterans. *Military Psychology, 18*(2), 89–120.

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics, 50*(6), 613–621.

Litwin, M. S. (2003). *How to assess and interpret survey psychometrics* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The construct of resilience: a critical evaluation and guidelines for future work. *Child Development, 71*(3), 543–562. doi:10.1111/1467-8624.00164

Maloney, P. W., Grawitch, M. J., & Barber, L. K. (2012). The multi-factor structure of the Brief Self-Control Scale: discriminant validity of restraint and impulsivity. *Journal of Research in Personality, 46*, 111–115. doi:10.1016/j.jrp.2011.10.001

Masten, A. S., Best, K. M., & Garmezy, N. (1990). Resilience and development: contributions from the study of children who overcome adversity. *Development and Psychopathology, 2*(4), 425–444. doi:10.1017/S0954579400005812

Meredith, L. S., Sherbourne, C. D., Gaillot, S., Hansell, L., Ritschard, H. V., Parker, A. M., & Wrenn, G. (2011). *Promoting psychological resilience in the U.S. military.* Santa Monica, CA: RAND Corporation. Retrieved from https://www.rand.org/pubs/monographs/MG996.html

Morin, C. M., Belleville, G., Bélanger, L., & Ivers, H. (2011). The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep, 34*(5), 601–608.

Naval Center for Combat and Operational Stress Control. (2016). *Servicemember Evaluation Tool [Fact Sheet]*.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy (ODASD (MC&FP)). (2017). *2015 demographics: profile of the military community.* Department of Defense. Retrieved from http://download.militaryonesource.mil/12038/MOS/Reports/2015-Demographics-Report.pdf

Online Etymology Dictionary. (n.d.). Retrieved October 12, 2017, from Online Etymology Dictionary: https://www.etymonline.com/word/resilience

Pavot, W., & Diener, E. (1993). Review of the Satisfaction with Life Scale. *Psychological Assessment, 5*(2), 164–172. doi:10.1207/s15327752jpa4901_13

Pavot, W., Diener, E., Colvin, C. R., & Sandvik, E. (1991). Further validation of the Satisfaction With Life Scale: evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment, 57*(1), 149–161. doi:10.1207/s15327752jpa5701_17

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research, 21*(2), 381–391.

Prins, A., Ouimette, P., Kimerling, R., Cameron, R. P., Hugelshofer, D. S., Shaw-Hegwer, J., . . . Sheikh, J. I. (2003). The primary care PTSD screen (PC-PTSD): development and operating characteristics. *Primary Care Psychiatry, 9*(1), 8–14. doi:10.1185/135525703125002360

Rice, V., & Liu, B. (2016). Personal resilience and coping part II: identifying resilience and coping among U.S. military service members and veterans with implications for work. *Work, 54*, 335–350. doi:10.3233/WOR-162301

Richardson, G. E., Neiger, B. L., Jensen, S., & Kumpfer, K. L. (1990). The resiliency model. *Journal of Health Education, 21*, 33–39.

Savard, M.-H., Savard, J., Simard, S., & Ivars, H. (2005). Emperical validation of the Insomnia Severity Index in cancer patients. *Psycho-Oncology*, 429–441. doi:10.1002/pon.860

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4*(3), 219–247.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67*(6), 1063–1078.

Secretary of the Navy. (2013). *Department of the Navy 21st Century Sailor and Marine initiative (SECNAVINST 5300.40).* Washington, DC: Department of the Navy.

Shattuck, N. L., & Matsangas, P. (2016). Comparison of the 3/9 and 6/6 watchstanding schedules for crewmembers of a U.S. Navy destroyer. *Proceedings of the Human Factors and Ergonomics Society 2016 Annual Meeting*, (pp. 881–885). doi:10.1177/1541931213601201

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271–322.

Vogt, D., Smith, B. N., King, L. A., King, D. W., Knight, J., & Vasterling, J. J. (2013). Deployment Risk and Resilience Inventory-2 (DRRI-2): an updated tool for assessing psychosocial risk and resilience factors among service members and veterans. *Journal of Traumatic Stress, 26*, 710–717. doi:10.1002/jts.21868

Ware Jr., J. E., & Sherbourne, C. D. (1992). The MOS 36-item Short-Form Health Survey (SF-36) I. Conceptual framework and item selection. *Medical Care, 30*(6), 473–483.

Wells, A., & Davies, M. I. (1994). The Thought Control Questionnaire: a measure of individual differences in the control of unwanted thoughts. *Behaviour Research and Therapy, 32*(8), 871–878.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California