NPS-OR-18-004

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

ASSESSMENT OF THE SAILOR EVALUATION TOOL (SET)

by

Nita Lewis Shattuck and Panagiotis Matsangas

November 2018

Approved for public release; distribution is unlimited

Prepared for: OPNAV/N1

THIS PAGE INTENTIONALLY LEFT BLANK

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

1. REPORT DATE ***(DD-MM-YYYY)***	**2. REPORT TYPE**	**3. DATES COVERED** ***(From-To)***
11-30-2018	Technical Report	October 2016 – November 2018

4. TITLE AND SUBTITLE
Assessment of the Sailor Evaluation Tool (SET)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Nita Lewis Shattuck and Panagiotis Matsangas

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Operations Research Department, Naval Postgraduate School; Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-OR-18-004

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
OPNAV/N1 Manpower, Personnel, Training and Organization

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES
The views expressed in this report are those of the author(s) and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

14. ABSTRACT
Psychological resilience is considered an essential aspect for building and maintaining operational readiness. To improve psychological resilience, appropriate tools are needed which are able to accurately measure psychological readiness. The Naval Center for Combat & Operational Stress Control (NCCOSC) developed the Servicemember Evaluation Tool (SET), a self-report battery that measures psychological resilience and well-being across several domains. We focused on crewmembers working on Unites States Navy ships while underway for a qualitative and quantitative assessment of the SET instrument. This effort included the utility of SET for data collection in the field, its psychometric properties, and the association between SET and other variables of interest, e.g., demographics, sleep attributes, and watchbill types. Data from 401 Sailors serving on four ships were collected during two ~1-week periods, at the beginning and close to the end of deployment. We concluded that although an instrument like the SET is a much-needed instrument to assess psychological resilience in active duty service members, the instrument in its current form should be revised in order to refine its focus, improve its psychometric properties and reduce potential sources of bias. Once these tasks are accomplished, a new data collection effort on multiple naval vessels could be used to reassess its utility.

15. SUBJECT TERMS
Resilience, Sleep-related behaviors, Sleep, Fatigue

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			
Unclassified	Unclassified	Unclassified	Unclassified	82	Nita Lewis Shattuck
					19b. TELEPHONE NUMBER ***(include area code)*** (831) 656-3481

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

The report entitled "Assessment of the Sailor Evaluation Tool (SET)" was prepared for OPNAV/N1 and funded by OPNAV/N171.

Further distribution of all or part of this report is authorized.

This report was prepared by:

Nita Lewis Shattuck, Ph.D.
Professor of Operations Research

Panagiotis Matsangas, Ph.D.
NRC Postdoctoral Associate

Reviewed by:

Johannes O. Royset
Associate Chairman for Research
Department of Operations Research

Released by:

Matt Carlyle
Chairman
Department of Operations Research

Jeffrey D. Paduan
Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Psychological resilience is considered an essential aspect for building and maintaining operational readiness. To improve psychological resilience, appropriate tools are needed which are able to accurately measure psychological readiness. The Naval Center for Combat & Operational Stress Control (NCCOSC) developed the Servicemember Evaluation Tool (SET), a self-report battery that measures psychological resilience and well-being across several domains. We focused on crewmembers working on Unites States Navy ships while underway for a qualitative and quantitative assessment of the SET instrument. This effort included the utility of SET for data collection in the field, its psychometric properties, and the association between SET and other variables of interest, e.g., demographics, sleep attributes, and watchbill types. Data from 401 Sailors serving on four ships were collected during two ~1-week periods, at the beginning and close to the end of deployment. We concluded that although an instrument like the SET is a much-needed instrument to assess psychological resilience in active duty service members, the instrument in its current form should be revised in order to refine its focus, improve its psychometric properties and reduce potential sources of bias. Once these tasks are accomplished, a new data collection effort on multiple naval vessels could be used to reassess its utility.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

A series of studies at the Naval Postgraduate School has demonstrated that fatigue due to sleep deprivation and circadian misalignment play an important role in the performance and morale of US Navy crews (Brown, Matsangas, & Shattuck, 2015a; Shattuck & Matsangas, 2014, 2015a, 2015b; Shattuck, Matsangas, & Brown, 2015; Shattuck, Matsangas, Moore, & Wegemann, 2015; Shattuck, Matsangas, & Powley, 2015). Scientific evidence points to a strong causal link between sleep and a variety of outcomes such as rate of wound healing and pain tolerance (Adam & Oswald, 1984; Christian, Graham, Padgett, Glaser, & Kiecolt-Glaser, 2006; Guo & DiPietro, 2010; Lautenbacher, Kundermann, & Krieg, 2006). Increased musculoskeletal pain has been associated with decreased sleep in the crew of a US Navy aircraft carrier (Brown, Matsangas, & Shattuck, 2015b; Shattuck, Matsangas, Moore, & Wegemann, 2016). In particular, working conditions – such as long hours, rotating shift work and night work – pose additional stress, which may lead to inadequate sleep and can be a risk factor for chronic disease (Shattuck, Matsangas, & Dahlman, 2018; Shattuck, Matsangas, Mysliwiec, & Creamer, 2019).

There are strong associations among sleep disorders and occupational factors that are common in military environments (e.g., rotating shiftwork, overseas deployments, chronic sleep restriction, and poor sleep quality) (Mysliwiec, Matsangas, Baxter, & Shattuck, 2015; Shattuck et al., 2019). These occupational factors may negatively influence health and contribute to the development of sleep disorders by disrupting normal sleep, adversely affecting circadian biological rhythms (both short- and long-term) and ultimately resulting in degraded safety, performance, and productivity of military populations. In addition, analysis of Defense Equal Opportunity Climate Survey (DEOCS) data indicates a strong negative correlation between amount of sleep and amount of perceived stress, which is known to degrade resilience over time.

According to the American Psychological Association (APA) "resilience is the process of adapting well in the face of adversity, trauma, tragedy, threats or significant sources of stress — such as family and relationship problems, serious health problems or workplace and financial stressors. It means "bouncing back" from difficult experiences." (American Psychological Association, 2018). The term also includes active resistance as well as recovery. The capacity for resilience in response to stressors is at least partially dependent on a well-functioning circadian system. Healthy sleep behaviors are associated with enhanced resilience and the ability to manage current and future stressors. Research on sleep has begun to shift from a focus on pathology and dysfunction to learning more about how healthy sleep promotes physical and psychological health and well-being. Sleep impacts overall stress levels and various areas of resilience such as mood (e.g., optimism and happiness), coping, cognitive flexibility, frustration tolerance, as well as behavioral and cognitive control (Pedersen et al., 2015). Impaired cognitive functioning often results in poor decisions that, for the warfighter, can have consequences such as costly damage to systems, potential injury or loss of life, and failure to accomplish the mission.

Multiple studies have shown that the military occupation is characterized by sleep deprivation and elevated fatigue levels with all three components of good sleep (timing, duration, and quality) being challenged in the operational environment (Miller, Matsangas, & Kenney, 2012; Miller, Matsangas, & Shattuck, 2008; Shattuck et al., 2018; Shattuck et al., 2019; Troxel et al., 2015). Given the amount of stress experienced by military personnel and the prevalence of sleep problems in the military, critical avenues of inquiry are how sleep impacts operational readiness and performance and how sleep can be maintained and improved in the context of operational demands. To date, there remains a gap in the development and evaluation of programs and well-controlled studies to promote healthy sleep patterns amongst deployed service members (Pedersen et al., 2015).

Along these lines, psychological resilience is considered an essential aspect to building and maintaining operational readiness. To improve psychological readiness, however, tools are needed which can measure psychological resilience. With this need in mind, subject-matter experts at the Naval Center for Combat & Operational Stress Control (NCCOSC) developed the Servicemember Evaluation Tool (SET), a self-report battery that measures psychological resilience and well-being across several domains. According to NCCOSC, SET can be helpful at both the individual level and the unit level in order to identify areas that individuals/units are doing well in and areas that need to be strengthened.

B. STUDY AIM AND GOALS

This study has the following objectives:

- To assess the relationship between individual resilience as measured by the Servicemember Evaluation Tool (SET) and the work and rest schedules Sailors experience during underway operations.
- To assess the relationship between sleep and overall Sailor resilience as measured by the SET.
- To assess ship resilience calculated as the sum total of the crewmembers' resilience scores.
- To assess the impact of watch schedules (and associated sleep) on behavioral factors (destructive behaviors, etc.).

THIS PAGE INTENTIONALLY LEFT BLANK

II. METHODS

A. EXPERIMENTAL DESIGN

All data collections were longitudinal in nature and were on based on a quasi-experimental approach: a naturalistic observation of U.S. Navy Sailors performing their normal underway duties while deployed on various surface combatants.

B. PARTICIPANTS

Recruited participants (N=401) were volunteers from three Arleigh Burke-class destroyers (USS CHAFEE – DDG 90, USS PINCKNEY – DDG 91, USS KIDD – DDG 100), and one Ticonderoga-class cruiser (USS PRINCETON – CG 59). Due to missing responses on their SET questionnairess, 41 Sailors were omitted from further analysis. Therefore, analysis was based on 360 Sailors, 235 who participated only in the Phase 1 data collection, 64 who participated only in Phase 2, and 61 who participated in both Phases. From the 360 participants, 142 were on USS CHAFEE, 95 on USS PRINCETON, 94 on USS KIDD, and 29 on USS PINCKNEY. Figure 1 shows the number of participants by study phase, and Table 1 shows participants' demographic information (N = 360) at the beginning of their participation.

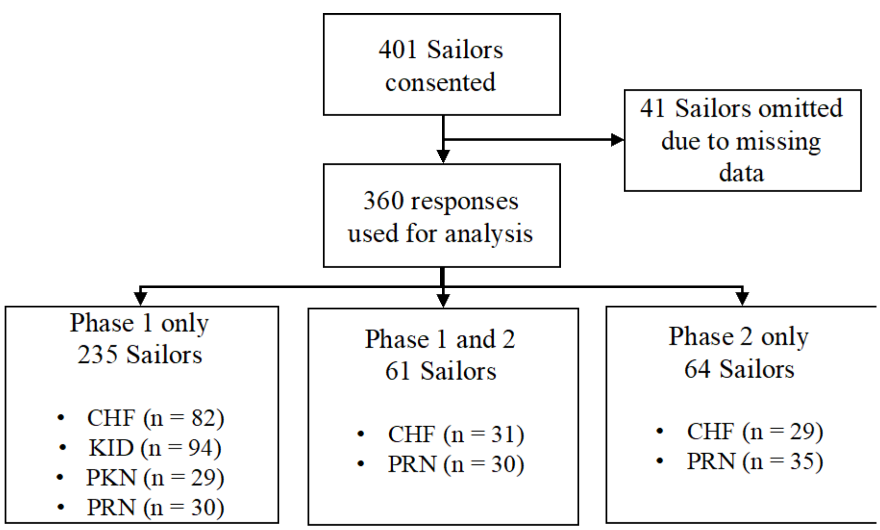


Figure 1. Study participants

Table 1. Demographic information.

Demographics	Entire sample (n = 360)
Age (years), MD ± MAD	27 ± 4
Gender, Males, # (%)	279 (77.5%)
Rank, # (%)	
Officers	62 (17.2%)
Enlisted	298 (82.8%)
Active duty (years), MD ± MAD	4.0 ± 2.5
Times deployed, MD ± MAD	1 ± 1
Total time deployed (months), MD ± MAD	7 ± 6
Department, # (%)	
Air	10 (2.78%)
Combat Systems	59 (16.4%)
Engineering	66 (18.3%)
Executive	31 (10.3%)
Operations	79 (21.9%)
Plans and Tactics	31 (8.61 %)
Supply	42 (11.7%)
Weapons	36 (10.0%)
Morningness-Eveningness (ME) Preference	
ME Score, M ± SD	54.1 ± 8.39
ME type, # (%)	
Definitely morning	10 (2.79%)
Moderately morning	93 (26.0%)
Intermediate	232 (64.8%)
Moderately evening	21 (5.87%)
Definitely evening	2 (0.56%)

C. EQUIPMENT AND INSTRUMENTS

1. Surveys

The self-administered morningness-eveningness questionnaire (MEQ-SA) (Terman, Rifkin, Jacobs, & White, 2001) was used to assess participants' chronotype, an attribute of humans reflecting their preference for waking earlier or later in the day. The scale includes 19 multiple-choice questions. Scores range from 16 to 86, with scores less than 42 corresponding to evening chronotypes and scores higher than 58 indicating morning chronotypes. Although based on the Horne and Östberg (1976) original MEQ scale, the MQE-SA has some stem questions and item choices rephrased to conform with spoken American english. Discrete item choices have been substituted for continuous graphic scales.

The ESS was used to assess average daytime sleepiness (Johns, 1991). The individual uses a 4-item Likert scale to rate the chance of dozing off or falling asleep in eight different everyday situations. Scoring of the answers was 0 to 3, with 0 being “would never doze,” 1 being “slight chance of dozing,” 2 being “moderate chance of dozing,” and 3 denoting a “high chance of dozing.” Respondents were instructed to rate each item according to his/her usual way of life in recent times. Responses were summed to the total score. A score of 10 or more reflects above normal daytime sleepiness and a need for further evaluation (Johns, 1992). The ESS questionnaire has a high level of internal consistency, as measured by Cronbach’s alpha, ranging from 0.73 to 0.88 (Johns, 1992). The 7-item Insomnia Severity Index (ISI) was used to assess the severity of both nighttime and daytime components of insomnia (Bastien, Vallieres, & Morin, 2001; Morin, Belleville, Bélanger, & Ivers, 2011).

The prestudy questionnaire included demographic information and one standardized questionnaire, the MEQ-SA. Demographic questions included age, gender, rate/rank, department, years on active duty, total months deployed, factors affecting sleep, type and frequency of caffeinated beverage use (i.e., tea, coffee, soft drinks, energy drinks), type and frequency of tobacco product use (i.e., cigarettes, chewing tobacco, nicotine gum or patches, electronic smoke), use of medication (prescribed or over-the-counter), and the type and frequency of any exercise routines.

The posttest questionnaire included the ESS, ISI, PSQI, and the Servicemember Evaluation Tool (SET). Participants were asked to indicate their watchstanding schedule, the adequacy of their own and their peers’ sleep (5-point Likert scale: “Much less than needed”; “Less than needed”; “About right”; “More than needed”; “Much more than needed”), and to compare their workload during the data collection period with their normal workload underway (5-point Likert scale: “Much less than usual,” “Less than usual”; “About the same”; “More than usual”; “Much more than usual”). The posttest questionnaire also included two open-ended questions (“What did you like most about your current watch schedule?” and “What did you like least about your current watch schedule?”).

a. Servicemember Evaluation Tool (SET)

The SET was developed at the Naval Center for Combat & Operational Stress Control (NCCOSC). According to NCCOSC, many resilience programs/models and assessment tools were reviewed, including the Department of Defense (DoD) Total Force Fitness (TFF) model (Bates et al., 2013), the Army's Comprehensive Soldier Fitness (CSF) program and Global Assessment Tool (GAT) (Casey, 2011; Peterson, Park, & Castro, 2011), and the 2011 RAND report "Promoting Psychological Resilience in the U.S. Military" (Meredith et al., 2011). Measures for the SET were chosen among psychological questionnaires with evidence of reliability and validity and, whenever possible, validated in military populations. Conceptually, the design of SET assumed that individual resilience domains include optimism, positive coping, behavioral control, flexible thinking, control and confidence, effective coping with stressful situations, life satisfaction, social support, emotional affect, perceived stress and psychological distress.

The version of SET used in the study included 25 component scores derived from 17 separate scales, i.e., 11 validated questionnaires, 3 subscales from validated questionnaires, and 3 independent items. According to NCCOSC, the 18 scales assessed General Resilience, Individual Resilience Factors, and Other Factors Related to Resilience.

<u>General Resilience</u>

The 10-item Connor-Davidson Resilience Scale (CD-RISC) was used to assess the ability to withstand, recover, grow and function competently in the face of stressors, adversity and changing demands (Campbell-Sills & Stein, 2007). Each item is scored on a 5-point Likert scale from 'not true at all' (0) to 'true nearly all the time' (4). Responses are summed for a total score ranging from 0 to 40. Higher scores are associated with greater resilience (Windle, Bennett, & Noyes, 2011).

The 4-item Response to Stressful Experiences Scale (RSES-4) was used to assess cognitive, emotional, and behavioral responses to stressful situations (De La Rosa, Webb-Murphy, & Johnston, 2016). Each item is scored on a 5-point Likert scale (0 – 4) from ‘not at all like me’ (0) to ‘exactly like me’ (4). Responses are summed for a total score ranging from 0 to 16. Higher scores are associated with greater resilience.

Individual Resilience Factors

The 18-item Brief COPE (BC) assesses different types of coping styles (i.e., active coping, denial, using emotional support, behavioral disengagement, using instrumental support, positive reframing, self-blame, planning, religion) (Carver, 1997). Each item is scored from 0 (‘I haven't been doing this at all’) to 3 (‘I've been doing this a lot’). Each subscale score is comprised from summing two items for a subscale score ranging from 0 to 6. Each subscale is interpreted separately, with subscale scores not summed for a total score. Higher scores are associated with greater use of a coping strategy. Of note, denial, behavioral disengagement, and self-blame are dysfunctional/negative coping strategies associated with worse resilience scores (Cooper, Katona, & Livingston, 2008; Rice & Liu, 2016).

The 8-item New General Self-Efficacy Scale (NGSES) was used to assess one’s perception of his/her ability to perform successfully across a variety of different situations (Chen, Gully, & Eden, 2001). Each item is scored on a 5-point Likert scale from ‘strongly disagree’ (1) to ‘strongly agree’ (5). Responses are summed for a total score ranging from 8 to 40. Higher scores are associated with greater self-efficacy.

The 13-item Brief Self-Control Scale (BSS) was used to assess one's ability to exert self-control, including thoughts, emotions, impulses, and performance (Tangney, Baumeister, & Boone, 2004). Each item is scored on a 5-point Likert scale from ‘not at all’ (1) to ‘very much’ (5). Nine items are reverse scored (#2, #3, #4, #5, #7, #9, #10, #12, #13). Responses are summed for a total score ranging from -41 to 11. Higher scores are associated with greater self-control.

The 6-item Life Orientation Test – Revised (LOT-R) was used to assess one’s expectancies for positive versus negative outcomes (Scheier & Carver, 1985). Each item is scored on a 5-point Likert scale from ‘strongly disagree’ (0) to ‘strongly agree’ (4). Three items are reversed scored (2, 4, 5). Responses are summed for a total score ranging from -12 to 12. Higher scores are associated with greater optimism.

The 6-item reappraisal subscale of the Thought Control Questionnaire (TCQ) was used to assess the tendency to use cognitive reappraisal when faced with unpleasant/unwanted thoughts (Wells & Davies, 1994). Each item was scored on a 4-point Likert scale from ‘never’ (0) to ‘almost always’ (3). Responses were summed for a total score ranging from 0 to 18. Higher scores were associated with greater use of cognitive reappraisal.

The 4-item positive reappraisal subscale of the Cognitive Emotional Regulations Questionnaire (CERQ) was used to measure the extent that one can reinterpret negative/unpleasant events in a positive way (Garnefksi & Kraaij, 2007). Each item was scored on a 5-point Likert scale from ‘almost never’ (1) to ‘almost always’ (5). Responses were summed for a total score ranging from 4 to 20. Higher scores were associated with more frequent use of positive reappraisal.

Other Factors Related to Resilience

To measure mood states and assess changes in mood, participants filled out the POMS (McNair, Lorr, & Droppelman, 1971). The POMS is a standardized, 65-item inventory originally developed to assess mood state in psychiatric populations. The questionnaire assesses various dimensions of the mood construct using six subscales: anger - hostility (12 items; range 0-48), confusion - bewilderment (7 items; range 0-28), depression (15 items; range 0-60), fatigue (7 items; range 0-28), tension - anxiety (9 items; range 0-36) and vigor - activity (8 items; range 0-32). Vigor is subtracted and the Total Mood Disturbance (TMD) score is derived by adding the subscales (range -32 to 200). Normalized scores (T-scores) are based on norms for adults (Nyenhuis, Yamamoto, Luchetta, Terrien, & Parmentier, 1999). The POMS was administered using the instruction set: “Describe how you felt during the past two weeks.”

The 12-item Unit Support Scale (USS), derived from the Deployment Risk and Resilience Inventory-2 (DRRI-2), was used to assess the emotional closeness that military personnel perceive to have with other members (peers, leaders) of the military (Vogt et al., 2013). Each item was scored on a 5-point Likert scale from ‘strongly disagree’ (1) to ‘strongly agree’ (5). Responses were summed for a total score ranging from 12 to 60 points. Higher scores were associated with greater perceived social support from fellow unit members and unit leaders.

The 5-item Satisfaction with Life Scale (SWLS) was used to assesses one's satisfaction with life as a whole (Pavot & Diener, 1993). Each item was scored on a 7-point Likert scale from 'strongly disagree' (1) to 'agree strongly' (7). Responses were summed for a total score ranging from 5 to 35. Higher scores were associated with greater satisfaction with life.

The 4-item Perceived Stress Scale (PSS) was used to assess one's perception of how much stress he/she experienced over the past month (Cohen & Williamson, 1988). Each item was scored on a 5-point Likert scale from 'never' (0) to 'very often' (4). Two items were reversed scored (#2, #3). Responses were summed for a total score ranging from -8 to 8. Higher scores were associated with higher perceived stress.

The 4-item Primary Care Posttraumatic Stress Disorder (PC-PTSD) screen tool was used to assess PTSD symptoms (Prins et al., 2003). Items were scored with 1 point for 'yes,' and 0 points for 'no.' Responses were summed for a total score ranging from 0 to 4 points. Higher scores were associated with more severe PTSD symptoms.

The 4-item Patient Health Questionnaire - 4 (PHQ-4) was used to assessed depression and anxiety symptoms (Kroenke, Spitzer, Williams, & Löwe, 2009). Each item was scored on a 4-point Likert scale from 'not at all' (0) to 'nearly every day' (3). Responses were summed for a total score ranging from 0 to 12 points. Higher score indicated more severe symptoms of depression and anxiety.

Participants were asked to rate their personal morale (5-point Likert scale; 'very low' – 0 to 'very high' – 4), and the morale in their unit (5-point Likert scale; 'very low' – 0 to 'very high' – 4). Using the single item from SF-36, participants rated their health status with a 5-point Likert scale from 'poor' (0) to 'excellent' (4).

2. Sleep assessment

Sleep was assessed by wrist-worn actigraphy supported by activity logs, a validated method to collect objective sleep data in field studies (Meltzer, Walsh, Traylor,

& Westin, 2012; Rupp & Balkin, 2011). The use of actigraphy followed existing recommendations (Ancoli-Israel et al., 2015; Morgenthaler et al., 2007). Two types of actigraphs were used, the Motionlogger Watch (Ambulatory Monitoring, Inc. [AMI]; Ardsley, New York), and the Spectrum (Philips-Respironics [PR]; Bend, Oregon) actiwatch. Data for both devices were collected in 1-minute epochs. AMI data (collected in the Zero-Crossing Mode) were scored using Action W version 2.7.2155 software. The Cole-Kripke algorithm with rescoring rules was used. Criterion for sleep and wake episodes was five minutes. The sleep latency criterion was no more than one minute awake in a 20-minute period (all values are default for this software). PR data were scored using Actiware software version 6.0.0 (Phillips Respironics; Bend, Oregon). The medium sensitivity threshold (40 counts per epoch) was used, with 10 immobile minutes the criterion for sleep onset and sleep end (all values are default for this software). Previous research has shown that AMI data analyzed with Cole-Kripke and PR data analyzed with medium sensitivity parameters assess total sleep time for an approximately 8-hour night sleep episode with 3-minute precision (average results compared to polysomnography derived 436 minutes of sleep) (Meltzer et al., 2012). Furthermore, results from a study conducted with Sailors from the Reactor Department of the USS NIMITZ showed that daily duration of time in bed (TIB) or daily total sleep time (TST) did not differ between participants wearing the AMI and PR actiwatches (Wilcoxon Rank Sum test for all differences $p > 0.500$) (Shattuck & Matsangas, 2015b).

3. Activity Logs

All participants were asked to complete an activity log, documenting their daily routine in accordance with Navy Availability Factor (NAF) categories (OPNAV, 2015). The activity logs covered a 24-hour period in 15-minute intervals. Participants were asked to document in the log whether they were exposed to sunlight (duration, and timing), consumption of caffeinated beverages and energy drinks, and whether they worked out, including time and duration of workout.

4. Other sources of information

We also extracted information regarding sick call visits from the Military Health System Management Analysis and Reporting Tool (M2), Driving Under the Influence (DUI) and Alcohol-Related Incidents (ARI) information extracted from the Alcohol and Drug Management Information System (ADMITS), Suicide Related Behaviors (SRBs), and information from the Web-Enabled Safety System (WESS) on naval mishap/hazard reports.

D. PROCEDURES

The study protocol was approved by the Naval Postgraduate School Institutional Review Board (NPS.2017.0022). Data collection was conducted in two phases. Phase 1 occurred between June and July 2017 and Phase 2 occurred between November 2017 and January 2018.

Initially, Sailors were briefed on the research protocol and study procedures. Individuals who agreed to participate in the study signed informed consent forms and received further training prior to being issued study equipment. Sailors who decided to participate in the Study completed the Pre-study Questionnaire and received their actiwatches and activity logbooks. All participants were instructed to fill out their activity logs daily. The AMI actiwatches were administered to portions of the watchstanders. These participants were instructed to perform the PVT prior to and after their watchstanding period. Phillips-Respironics actiwatches were administered to the rest of the participants. Upon completion of the study, the participants returned their equipment and filled out an end-of-study questionnaire. These procedures were repeated twice, in the beginning and at the end of the deployment. Operational reasons did not allow for collecting data in the midpoint of the deployment. Figure 2 shows an overview of the study.

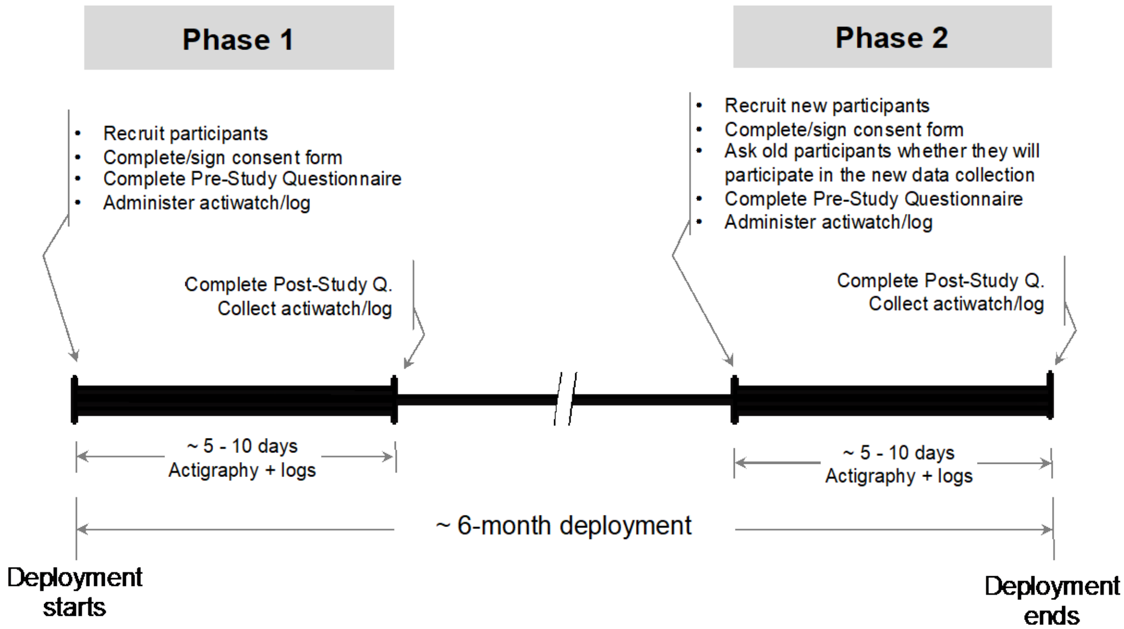


Figure 2. Study overview

During the first data collection, Sailors were working in various Watchstanding schedules, both circadian and non-circadian. Before the second data collection, however, a policy change was introduced in the USN ordering all surface combatants to use circadian schedules. Hence, during the second data collection, almost all Sailors who were working on shifts stood watch on circadian watchstanding schedules.

E. ANALYTICAL APPROACH

1. Actigraphy Data Cleaning and Reduction Procedures

The preparation of the actigraphy data for analysis included three steps. First, we evaluated the number of days of data available for each participant. Participants with fewer than five days of data were excluded from this analysis.

Next, we compared the actigraphy data with the activity logs. The primary source for the sleep analysis was the actigraphy data, but sleep logs assisted in the determination of start and end times of sleep intervals. Based on this comparison, we manually identified the start and end times of sleep episodes in the actigraphy data. The criteria used to determine whether we could use the data or whether imputation was required included the quality of the actigraphy data, the consistency of activity patterns over consecutive days, the amount of missing data, whether the participant was a watch-stander, and the accuracy of the sleep log. Imputation was applied only when: (a) there was a gap in actigraphy data within which the sleep log showed a sleep interval, and (b) the pattern of actigraphy data, assisted by the activity logs, was such to assure a confidence in the interpolation of a sleep interval.

Based on the actigraphy data, an initial database of sleep intervals was developed. From the rest/in-bed intervals (identified as DOWN in the AMI software, and REST in Phillips) the time in-bed (TIB) was calculated. Within each rest interval, the actigraphically assessed sleep was calculated.

2. Analysis Roadmap

Initially, all variables underwent descriptive statistical analysis to describe our population in terms of demographic characteristics, sleep-related behaviors, and scores in SET components. Next, we focus explicitly on SET assessment, both qualitatively and quantitatively. First, we assess basic psychometric properties of the SET using data from Phase 1. Correlation analysis was used to assess convergent and divergent (or discriminant) validity (Litwin, 2003, p. 42), i.e., avoid patterns of inter-relations among the SET components. To assess further the divergent, as well as the convergent, validity of the SET we calculated the average inter-component correlation. The seven components with inverse scoring (POMS TMD, BC-18 denial, BC-18 behavioral disengagement, BC-18 self-blame, PSS-4, PC-PTSD, and PHQ-4; higher scores denote less favorable attribute) were subtracted in the calculation. In general, most researchers agree that the average inter-item correlation for a set of items in a questionnaire should be between 0.20 and 0.40. When they are within this range, average inter-item correlation suggests that, even though the items are reasonably homogenous, they contain sufficiently unique variance so as to not be isomorphic with each other (Piedmont, 2014).

Exploratory factor analysis (EFA) was used to assess underlying latent constructs in the SET (DeVellis, 2003, p. 103; Fabrigar, Wegener, MacCallum, & Strahan, 1999; Henson & Roberts, 2006). We identified the number of factors to retain using Cattell's scree test (Cattell, 1966) and the Kaiser criterion, i.e., all factors have an eigenvalue greater than 1 (Kaiser, 1960). Maximum likelihood was used as the factoring method (Fabrigar et al., 1999), and promax (oblique) as the rotation method. Oblique rotations assume that the latent factors are correlated, an assumption appropriate when assessing human behavior (Costello & Osborne, 2005). Loadings of 0.40 or greater were used to interpret the results.

The next section focused on criterion validity of SET, that is, the association between SET and demographic/occupational variables. We performed a correlation analysis between SET components and variables of interest using Pearson's r. We also developed composite scores using both a unidimensional and a multidimensional approach. The calculation of the unidimensional composite score was based on the following algorithm. First, the score of each SET component was normalized from 0 to 1. The composite score was the average of the normalized component scores. The components with inverse scoring (POMS TMD, BC-18 denial, BC-18 behavioral disengagement, BC-18 self-blame, PSS-4, PC-PTSD, and PHQ-4; higher scores denote less favorable attribute) were subtracted in the calculation of the overall SET score. Hence, the composite SET score ranged from 0 to 1 with 1 denoting higher resilience. The multidimensional approach was based on using Principal Component Analysis (PCA) to reduce the dimensionality of the SET to four composite factors.

Multiple regression analysis was used to assess whether SET composite scores were associated with operationally relevant factors (e.g., characteristics of watchstanding schedules, sleep attributes). Models included watchbill type (fixed, rotating), number of sections in the watchbill, daily sleep duration, and number of sleep episodes per day as the potential predictor factors. Rank group (Officer, Enlisted) and ship were included as confounding factors.

Using the composite scores, we also assessed differences between ships. The ship scores were calculated as the sum of the corresponding Sailor scores. Using multiple regression, the first group of models included ships (A, B, C, D) as the potential predictor factor, whereas the ship groups (one ship using both circadian and noncircadian watchbills – "mixed," two ships using predominantly circadian watchbills – "circadian," one ship using predominantly noncircadian watchbills – "non-circadian") was the potential predictor factor in the second group of models. In both cases, rank (Officer, Enlisted) was a confounding factor.

The next step in the analysis was to assess differences between Sailors in the highest and the lowest quartile in terms of their unidimensional SET composite score. Analysis of Variance (ANOVA) and Fisher's Exact test were used for these pair-wise comparisons. Lastly, we assessed differences in SET components and SET composite scores between the two phases of the study. This comparison was based on the 61 Sailors from two ships who were in both Phase 1 and Phase 2. We used a mixed-effects model analysis to identify differences between study phases. Ship was included as a confounding factor, and study phase was the potential predictor factor.

Statistical analysis was conducted with a statistical software package (JMP Pro 14; SAS Institute; Cary, NC). Data normality was assessed for all study variables with the Shapiro-Wilk W test. Descriptive results are presented as mean (M) ± standard deviation (SD). Post-hoc statistical significance was assessed using the Benjamini–Hochberg False Discovery Rate (BH-FDR) controlling procedure (Benjamini & Hochberg, 1995). An overview of the analysis roadmap is shown in Figure 3.

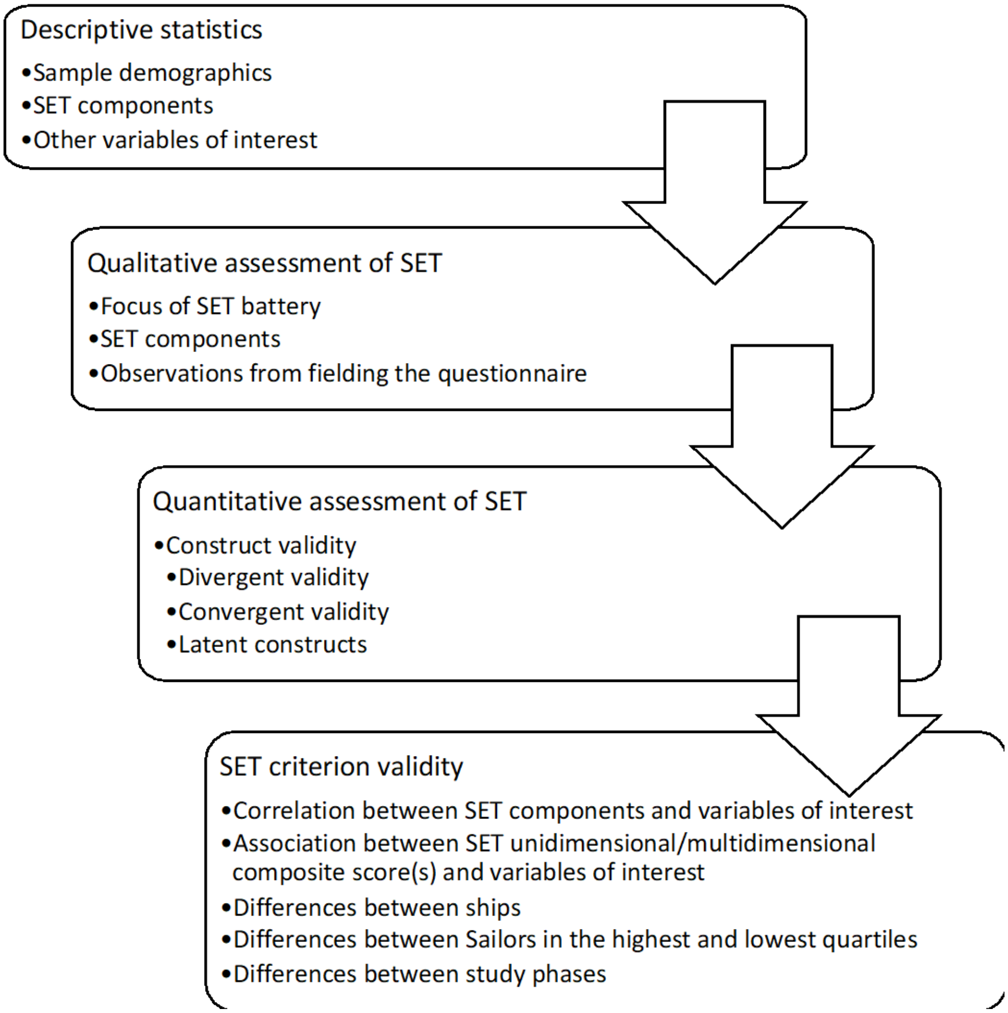


Figure 3. Analysis roadmap

III. RESULTS

A. DATA SAMPLE DESCRIPTION

1. Ship description in terms of watchbills used

Our analysis included data from four ships. During Phase 1 (the beginning of the underway), approximately 60% of the watchstanders on Ship A worked on non-circadian watchbills (predominantly 5hrs on/15hrs off) with the rest of the watchstanders working on circadian watchbills. This ship is identified as “mixed” because of the combination of both circadian and non-circadian watchbills. Most watchstanders (~70%) on ships B and C worked on circadian watchbills (predominantly 3hrs on/9hrs off), and those ships are identified as “circadian”. In contrast, most watchstanders on ship D worked on non-circadian watchbills (80% 5hrs on/15hrs off), and the ship is identified as “non-circadian”. This classification will be used later to assess differences in SET scores by ship.

2. Ship-wide metrics

Table 2 shows the number of DUI/ARI incidents, sick call visits, accidents/injuries, SRBs and mishaps/hazards reports between June and October 2017. During this period, for the most part, the ships were using the watchstanding schedule observed at the beginning of the deployment (Phase 1 of the data collection). Inspection of these data does not identify a recognizable pattern between the two circadian ships and the non-circadian ship.

Table 2. Ship-wide metrics

Ship	A (Mixed)	B (Circadian)	C (Circadian)	D (Non-circadian)
DUI incidents, #	0	0	0	2
ARI incidents, #	1	0	0	2
Sick call Visits, #	72	161	243	105
Sailors Fit for Full Duty, %	34.72%	87.58%	73.66%	66.67%
Sailors in Light Duty, #	0.00%	3.73%	16.05%	13.33%
Sailors Sick in Quarters, #	65.28%	8.70%	9.47%	20.00%
Sailors Medevac, #	0.00%	3.11%	1.23%	0.95%
Accidents and Injuries, #	5	11	8	9
Suicide Related Behaviors, #	0	5	0	0
Mishap/Hazard reports, #	4	8	0	11

3. Sleep duration

From the 360 Sailors in our study sample, 296 provided useful data in Phase 1 (261 Sailors had good actigraphic data). Analysis showed that crewmembers rested (that is, time spent in bed) on average, 7.48 ± 1.01 hours per day and slept for 6.55 ± 0.944 hours split in 1.43 ± 0.379 sleep episodes. Napping was frequent with 206 (78.9%) Sailors splitting their sleep in more than one sleep episode per day. Overall, 180 (69.0%) Sailors slept on average less than 7 hours per day, while 76 (29.1%) slept on average less than 6 hours per day.

4. Sleep-related attributes

Given these findings, it is no surprise that approximately 45% of the Sailors expressed negative opinions about the adequacy of their sleep. The same trend, but more pronounced (~62%), was found in their responses concerning the adequacy of the sleep of other Sailors (see Figure 4).

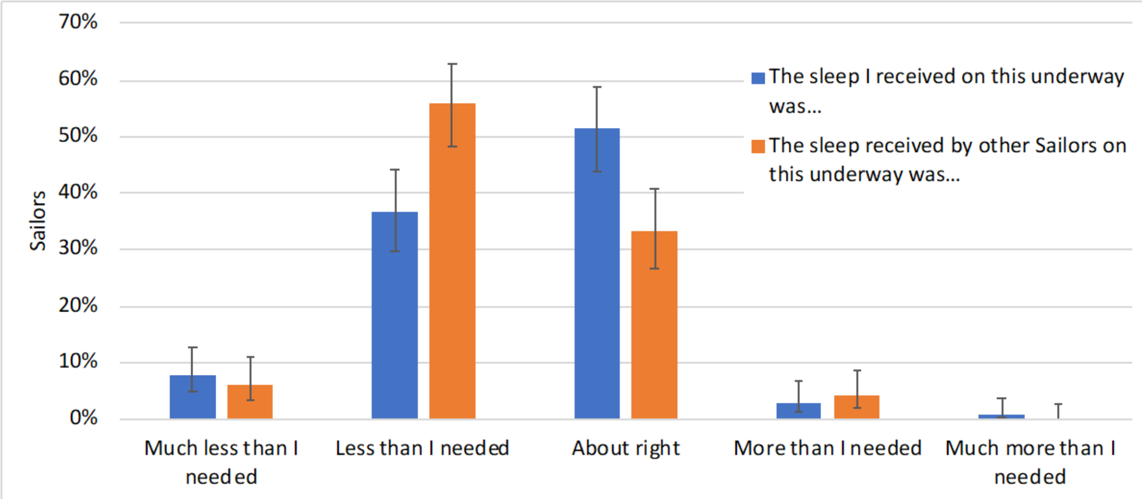


Figure 4. Responses to the statement "The sleep I received on this underway was…" and to the statement "The sleep received by other Sailors on this underway was…"

The study sample included 210 (71.0%) watchstanders and 86 (29.0%) non-watchstanders. Figure 5 shows detailed information regarding the watch/work schedule of the Sailors who participated in the study.

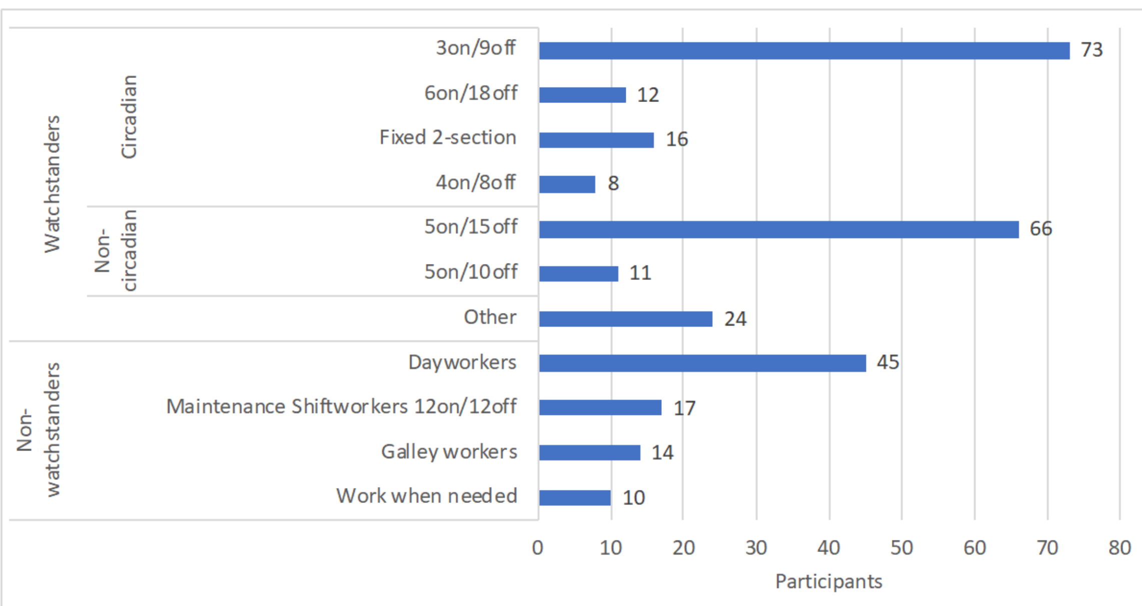


Figure 5. Watch/work schedule of the Sailors who participated in the study

The average ESS score was 10.3 ± 4.86 with 142 (48.0%) Sailors reporting elevated daytime sleepiness (ESS score > 10). The average ISI score was 10.9 ± 4.99

with 74 (25.0%) Sailors reporting elevated severity of insomnia symptoms (ISI score > 15). The average PSQI score was 8.23 ± 3.26 with 233 (78.7%) Sailors classified as poor sleepers (PSQI score > 5). In terms of POMS scores, the average POMS TMD score was 36.3 ± 34.5 (POMS T: 9.23 ± 6.29; POMS D: 11.2 ± 11.7; POMS A: 11.7 ± 9.33; POMS V: 13.3 ± 6.19; POMS F: 9.54 ± 6.32; POMS C: 7.99 ± 4.87). Compared to adult norms, more Sailors had worse mood in terms of TMD scores (65.2%, Likelihood Ratio test: $p < 0.001$), Anger-hostility (57.8%, $p = 0.007$), vigor-activity (77.7%, $p < 0.001$), fatigue (55.7%, $p = 0.048$), and confusion-bewilderment (63.9%, $p < 0.001$). These results are shown in Figure 6.

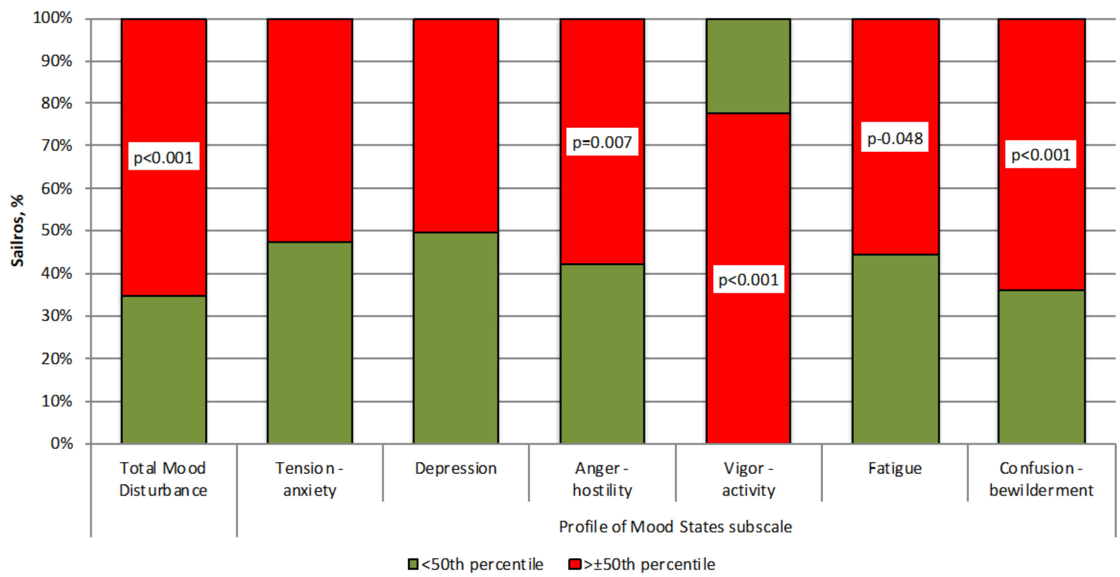


Figure 6. Percentage of Sailors with POMS scores worse than the 50th percentile of adult norms

1. Use of sleep-related substances

Participants reported the type and frequency of caffeinated beverages consumed (see Figure 5). Overall, 244 (82.7%) Sailors indicated drinking caffeinated beverages, with coffee being the most frequent (163, 55.3%), followed by soft drinks (93, 31.6%), energy drinks (85, 28.9%) and tea (57, 19.3%). For those Sailors who consumed these beverages, the median reported daily number of energy drinks consumed was 1 ± 0.5 cans, 1 ± 0.73 cans of soft drinks, 2 ± 1 cups of coffee, and 1 ± 0.5 cups of tea.

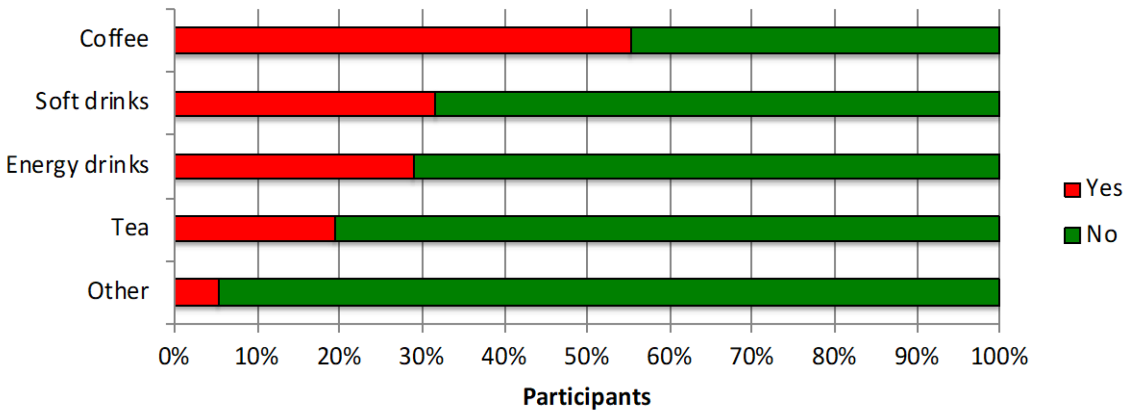


Figure 7. Consumption of caffeinated beverages.

Nicotine products were used by 96 (32.5%) participants, i.e., cigarettes (62, 21.0%), chewing tobacco/snuff (39, 13.2%), electronic smoke (5, 1.70%), cigar (n = 4), and Nicorette gum or nicotine patch (n = 1). Approximately two-thirds (208, 70.5%) of the participants reported having an exercise routine, working out on average 3 ± 3 times per week (median ± MAD), with a median duration of 1 ± 0.25 hour. The workout routines reported by the Sailors were mainly weight lifting and aerobic exercise.

2. SET components

The average score on the Connor-Davidson Resilience Scale (CD-RISC) was 28.1 ± 6.22 ranging from 6 to 40. Figure 2 shows the distribution plot of CD-RISK scores.

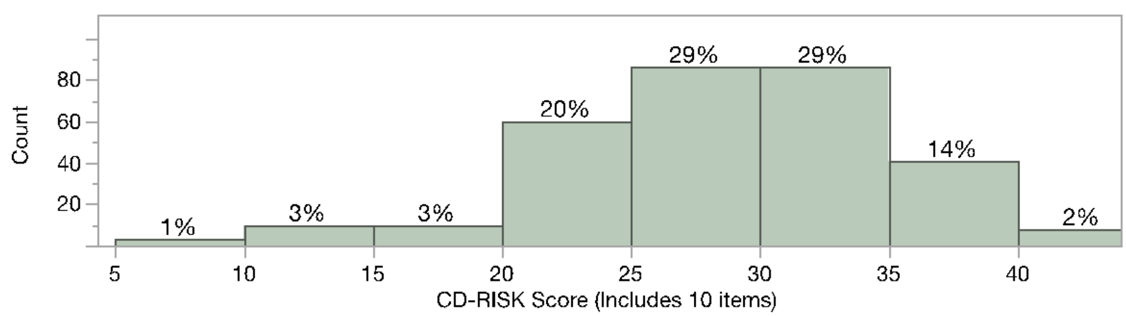


Figure 8. Distribution plot of CD-RISK scores

The average score on the Stressful Experiences Scale (RSES-4) was 12.1 ± 2.82 ranging from 1 to 16. Figure 3 shows the distribution plot of RSES-4 scores.

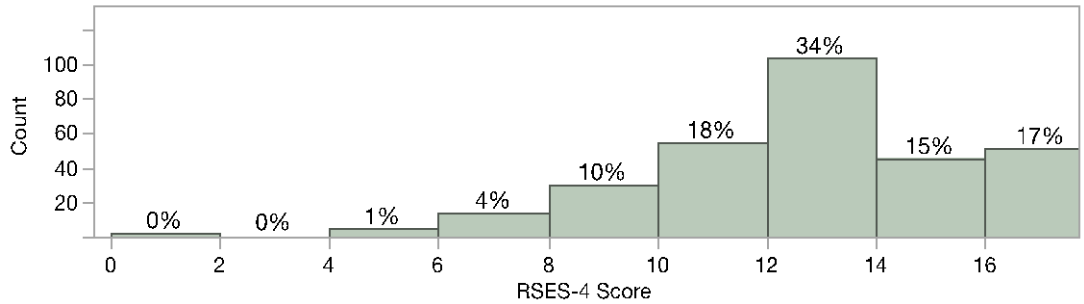
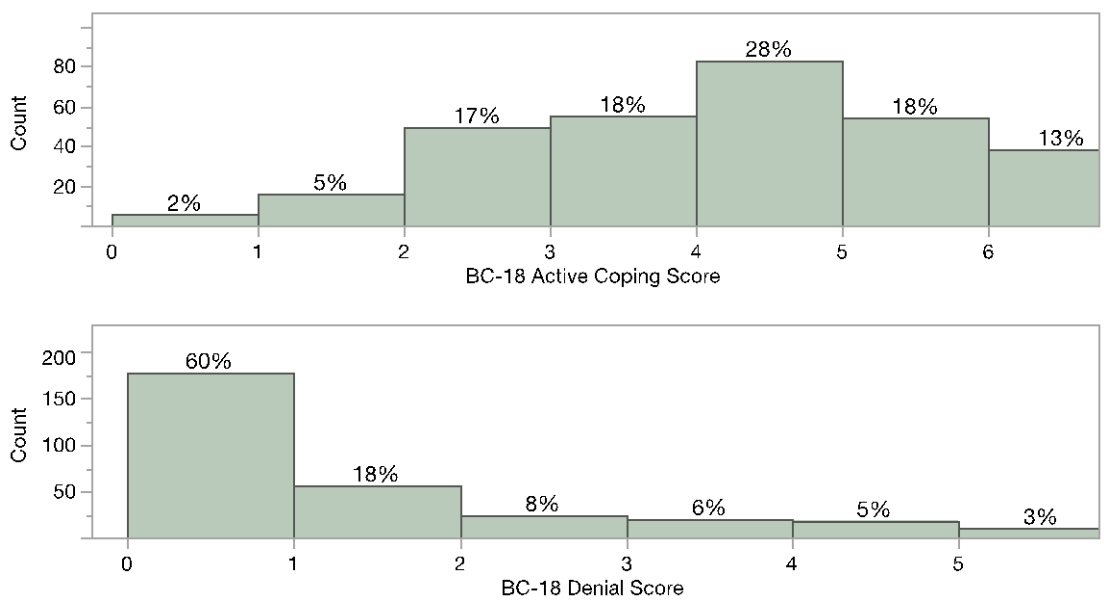
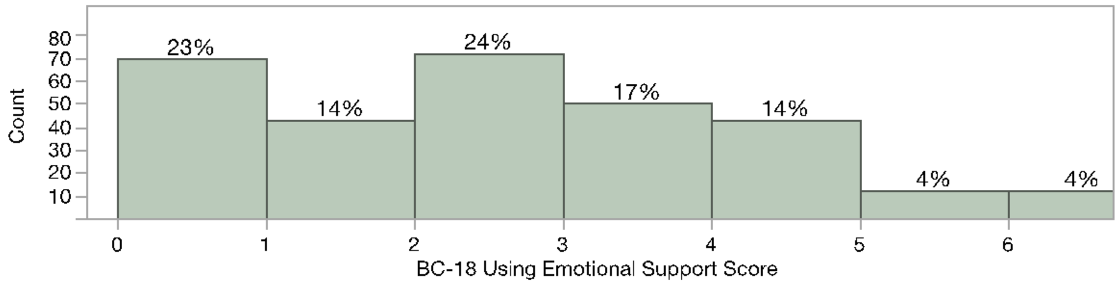
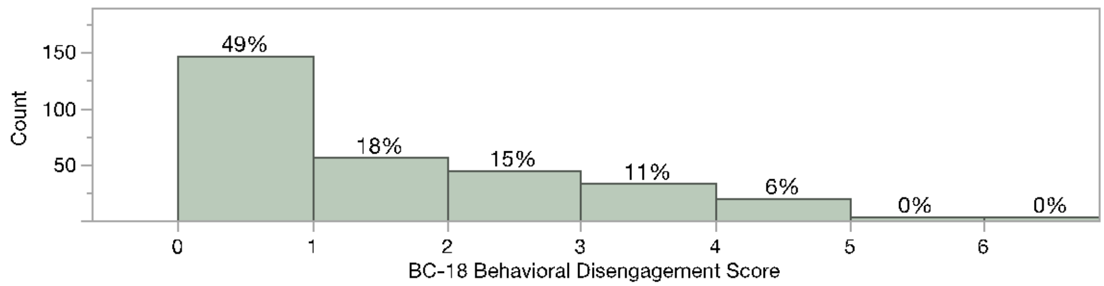


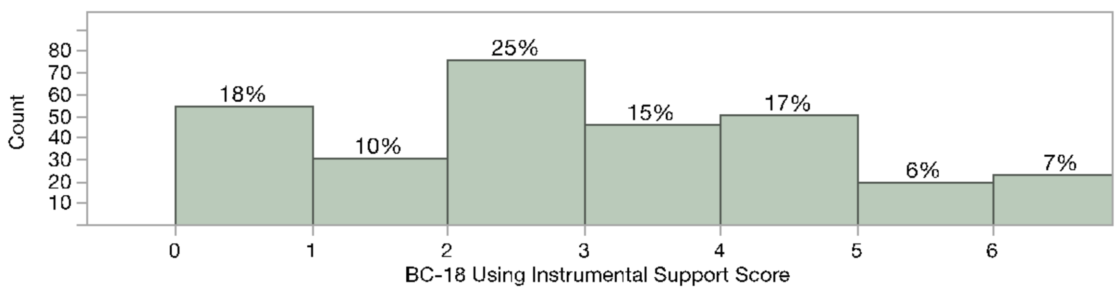
Figure 9. Distribution plot of RSES-4 scores

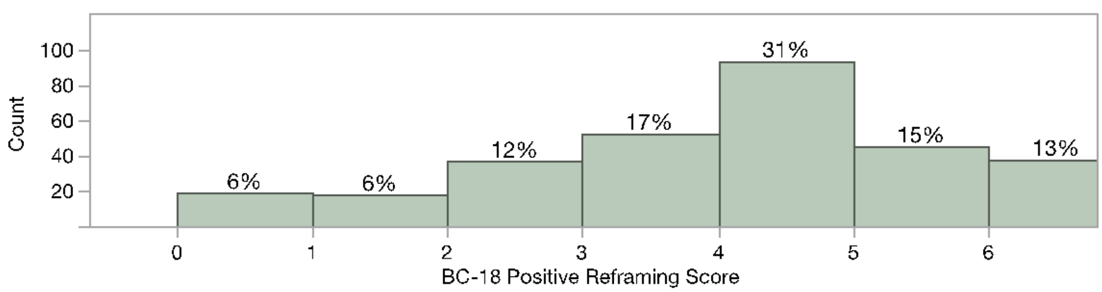
In terms of the BC-18 scales, the average score for active coping was 3.69 ± 1.47, 0.88 ± 1.35 for denial, 2.10 ± 1.65 for using emotional support, 1.10 ± 1.34 for behavioral disengagement, 2.51 ± 1.78 for using instrumental support, 3.56 ± 1.62 for positive reframing, 2.17 ± 1.72 for self-blame, 3.46 ± 1.65 for planning, and 1.86 ± 2.09 for religion. Scores on the Denial scale ranged from 0 to 5, but all other BC-18 scales ranged from 0 to 6. Figure 4 shows the distribution of BC-18 scale scores.

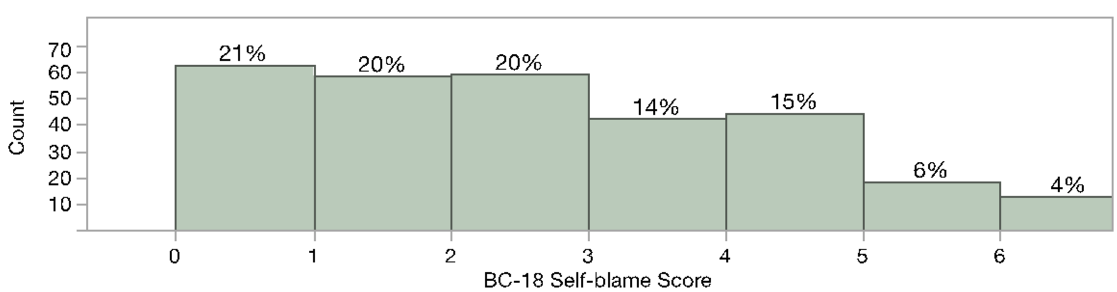



Count
80
70
60
50
40
30
20
10
23%
14%
24%
17%
14%
4%
4%
0
1
2
3
4
5
6
BC-18 Using Emotional Support Score


Count
150
100
50
49%
18%
15%
11%
6%
0%
0%
0
1
2
3
4
5
6
BC-18 Behavioral Disengagement Score


Count
80
70
60
50
40
30
20
10
18%
10%
25%
15%
17%
6%
7%
0
1
2
3
4
5
6
BC-18 Using Instrumental Support Score


Count
100
80
60
40
20
6%
6%
12%
17%
31%
15%
13%
0
1
2
3
4
5
6
BC-18 Positive Reframing Score


Count
70
60
50
40
30
20
10
21%
20%
20%
14%
15%
6%
4%
0
1
2
3
4
5
6
BC-18 Self-blame Score

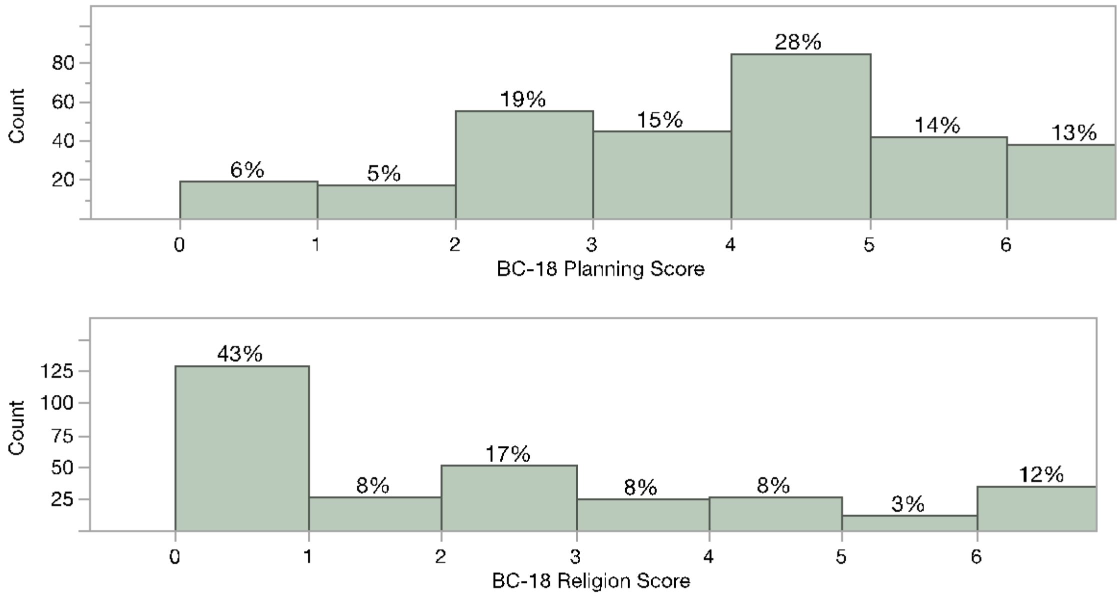


Figure 10. Distribution plots of BC-18 scores

The average score on the New General Self-Efficacy Scale (NGSES) was 32.5 ± 5.06 ranging from 8 to 40. Figure 5 shows the distribution plot of NGSES scores.

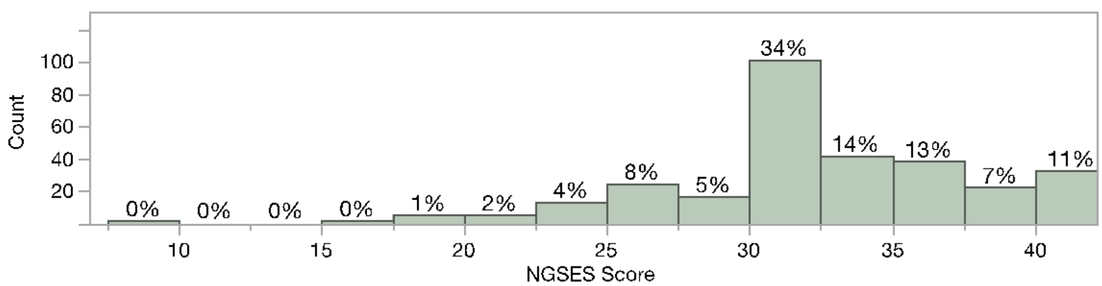


Figure 11. Distribution plot of NGSES scores

The average score on the Brief Self-Control Scale (BSS) was -12.0 ± 8.03 ranging from -33 to 11. Figure 6 shows the distribution plot of BSS scores.

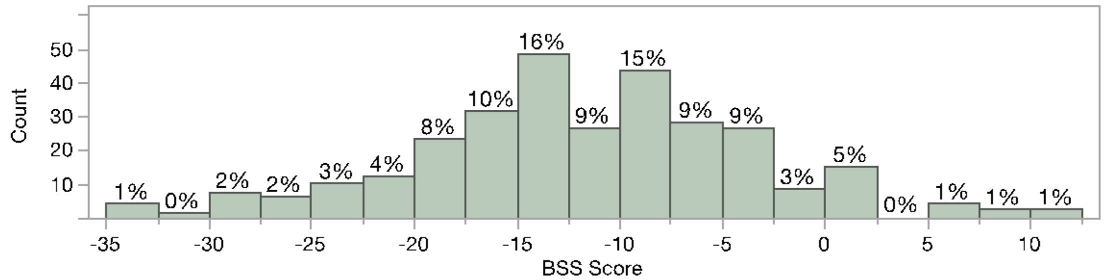


Figure 12. Distribution plot of BSS scores

The average score on the Life Orientation Test-Revised (LOT-R) was 1.97 ± 4.32 ranging from -12 to 12. Figure 7 shows the distribution plot of LOT-R scores.

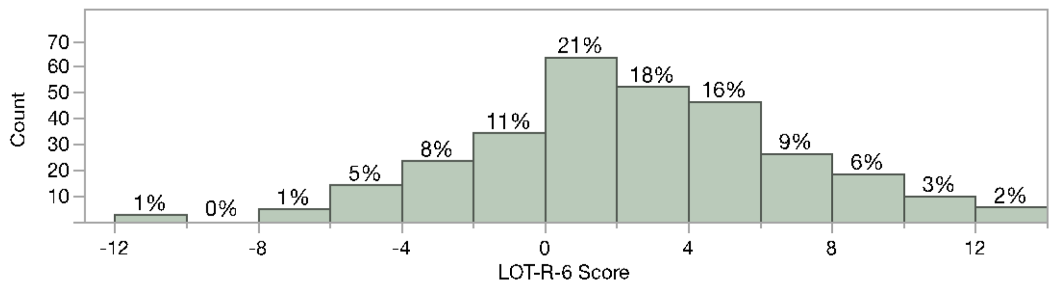


Figure 13. Distribution plot of LOT-R scores

The average score on the 6-item reappraisal subscale of the Thought Control Questionnaire (TCQ) was 9.86 ± 3.33 ranging from 0 to 18. Figure 8 shows the distribution plot of TCQ reappraisal subscale scores.

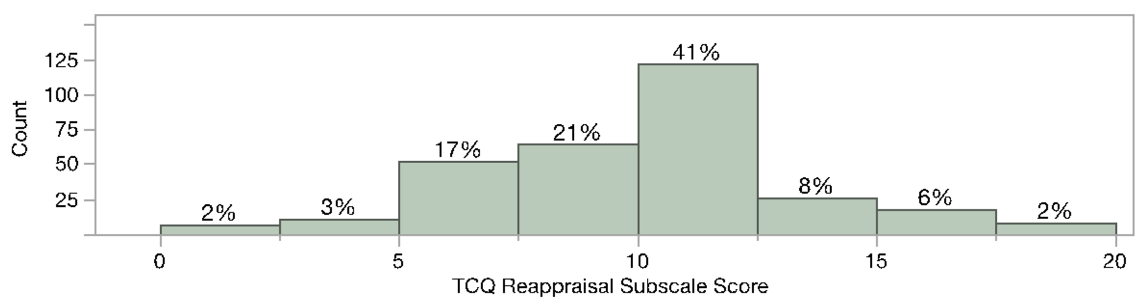


Figure 14. Distribution plot of the TCQ reappraisal subscale scores

The average score on the positive reappraisal subscale of the Cognitive Emotional Regulations Questionnaire (CERQ) was 14.1 ± 3.73 ranging from 4 to 20. Figure 9 shows the distribution plot of CERQ positive reappraisal subscale scores.

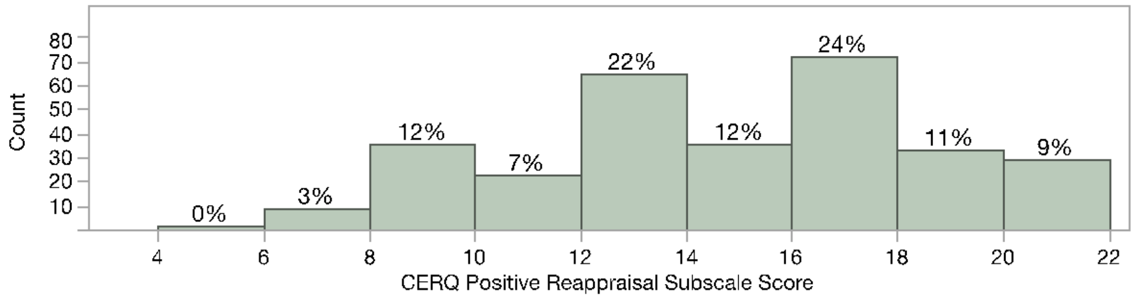
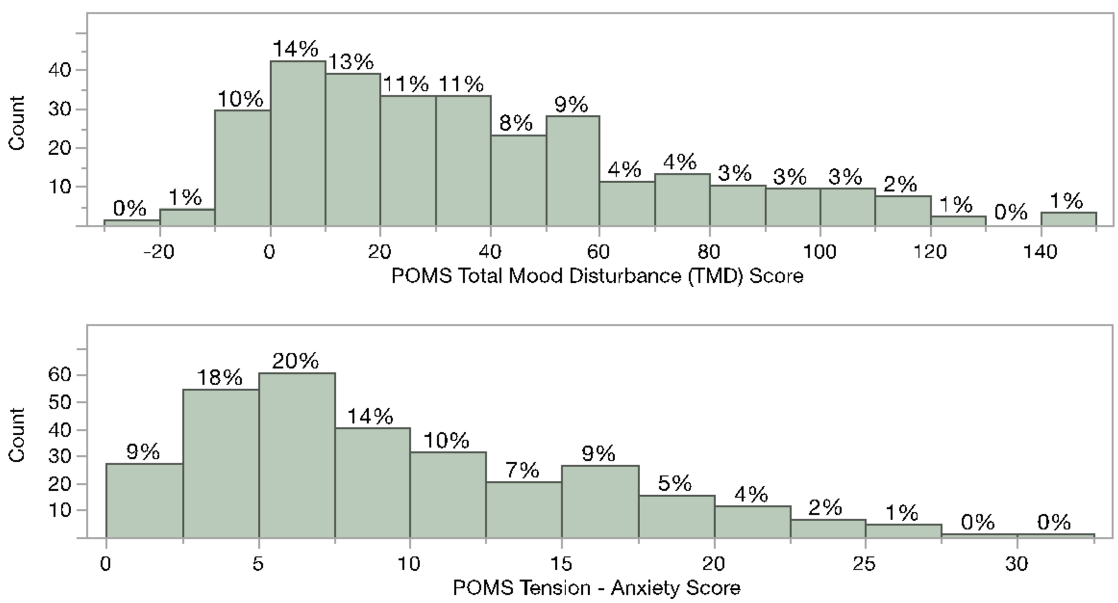


Figure 15. Distribution plot of CERQ positive reappraisal subscale scores

The average Total Mood Disturbance (TMD) score was 36.3 ± 34.5 ranging from -23 to 142. In terms of the POMS subscales, the average Tension/Anxiety score was 9.27 ± 6.29 ranging from 0 to 31, the average Depression score was 11.2 ± 11.7 ranging from 0 to 60, the average Anger/Hostility score was 11.7 ± 9.33 ranging from 0 to 44, the average Vigor score was 13.3 ± 6.19 ranging from 0 to 32, the average Fatigue score was 9.54 ± 6.32 ranging from 0 to 26, and the average Confusion/Bewilderment score was 7.99 ± 4.87 ranging from 0 to 23. Figure 10 shows the distribution of POMS scores.



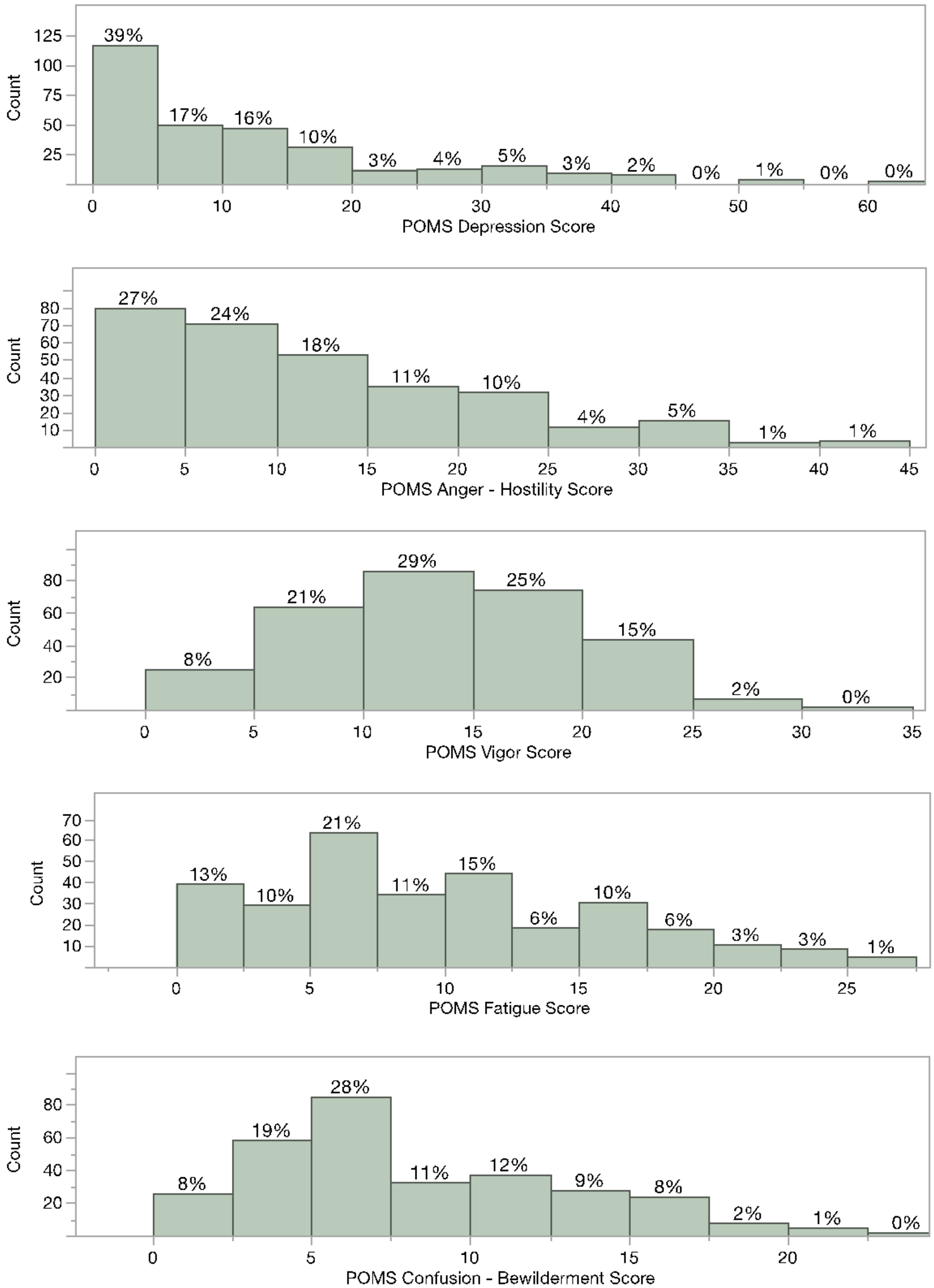


Figure 16. Distribution plots of POMS scores

Approximately 65% of the Sailors had TMD scores worse than the 50th percentile in adult norms (Likelihood Ratio test, $p < 0.001$), 52.4% of the Sailors had Tension/Anxiety scores worse than the 50th percentile (Likelihood Ratio test, $p = 0.416$), 50.3% had Depression scores worse than the 50th percentile (Likelihood Ratio test, p = 0.0.900), 57.8% of the Sailors had Anger/Hostility scores worse than the 50th percentile (Likelihood Ratio test, $p = 0.007$), 77.7% of the Sailors had Vigor scores worse than the 50th percentile (Likelihood Ratio test, $p < 0.001$), 55.7% of the Sailors had Fatigue scores worse than the 50th percentile (Likelihood Ratio test, $p = 0.048$), and 63.9% of the Sailors had Confusion/Bewilderment scores worse than the 50th percentile (Likelihood Ratio test, $p < 0.001$). These results are shown in the following Figure 17.

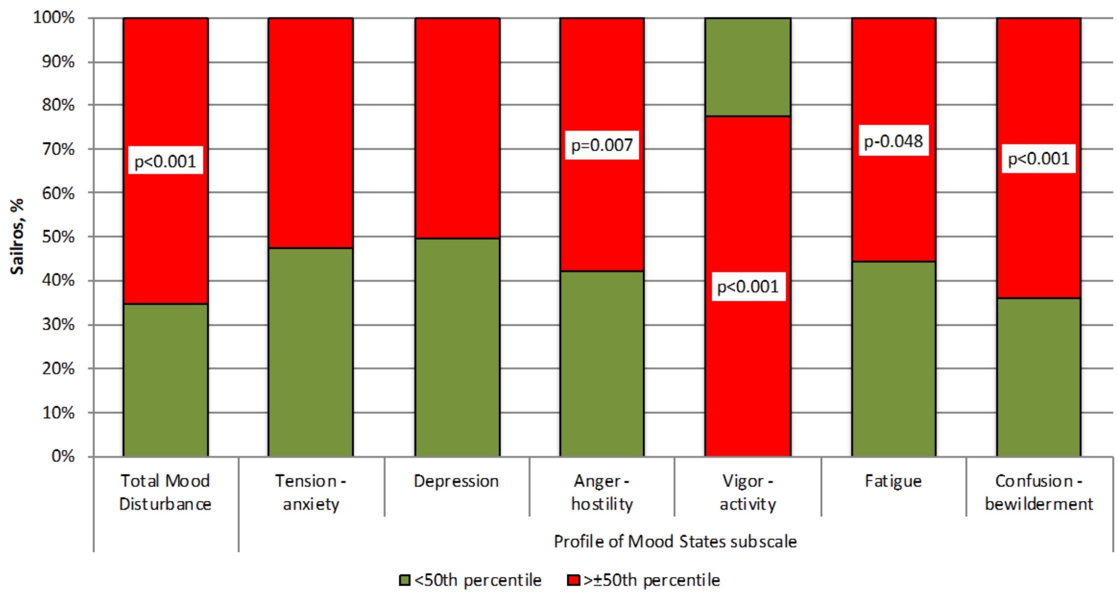


Figure 17. Percentage of Sailors with POMS scores worse the 50th percentile of adult norms

The average score on the Unit Support Scale (USS) was 40.2 ± 9.57, ranging from 12 to 60. Figure 12 shows the distribution plot of USS scores.

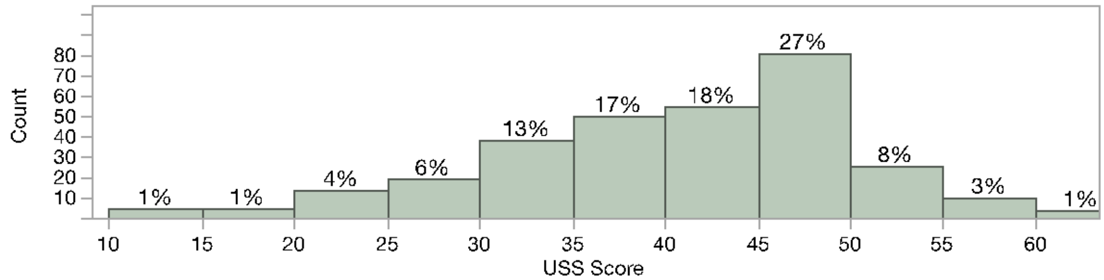


Figure 18. Distribution plot of USS scores

The average score on the Satisfaction with Life Scale (SWLS) was 22.2 ± 6.52, ranging from 5 to 35. Figure 13 shows the distribution plot of SWLS scores.

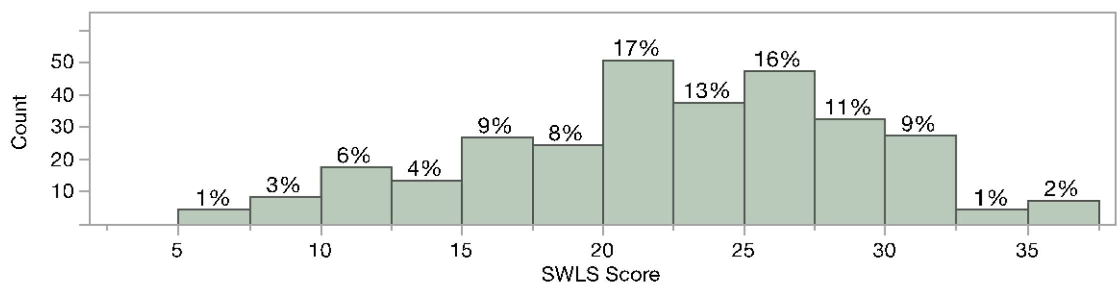


Figure 19. Distribution plot of SWLS scores

The average score on the Satisfaction with Perceived Stress Scale (PSS) was -1.17 ± 2.77 ranging from -8 to 7. Figure 14 shows the distribution plot of PSS scores.

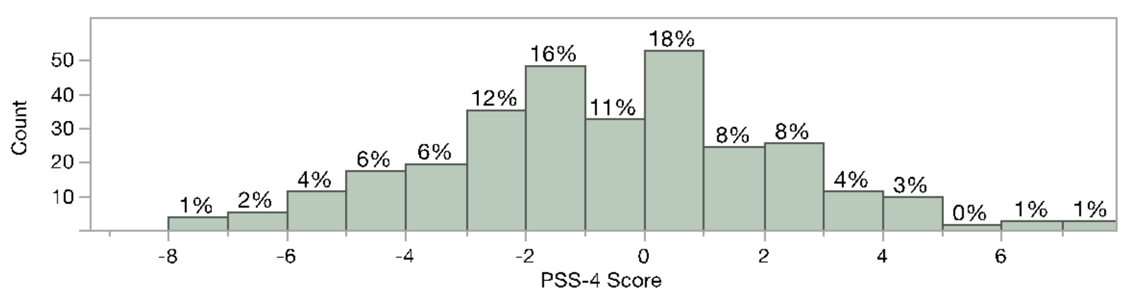


Figure 20. Distribution plot of PSS scores

The average score on the Primary Care Posttraumatic Stress Disorder (PC-PTSD) was 1.32 ± 1.46, ranging from 0 to 4. Figure 15 shows the distribution plot of PC-PTSD scores.

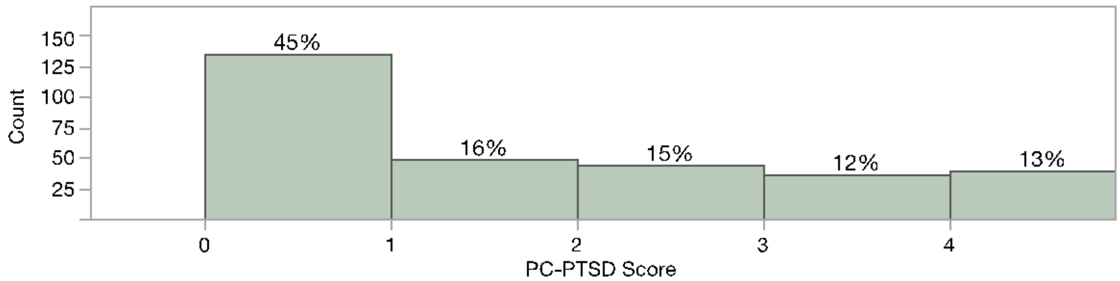


Figure 21. Distribution plot of PC-PTSD scores

The average score on the Patient Health Questionnaire - 4 (PHQ-4) was 2.94 ± 2.78, ranging from 0 to 12. Figure 16 shows the distribution plot of PHQ-4 scores.

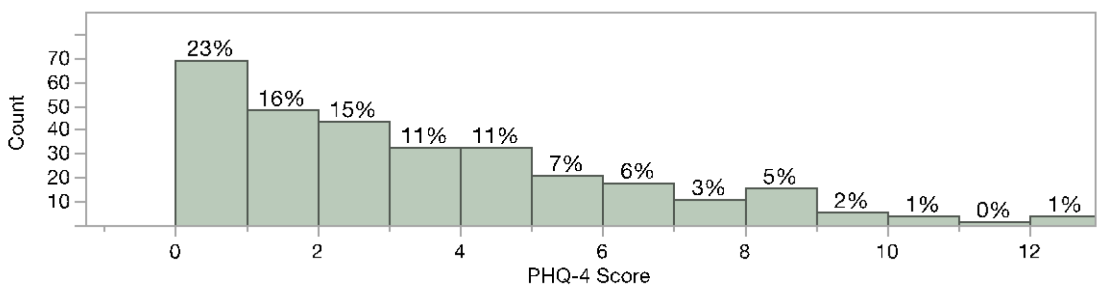


Figure 22. Distribution plot of PHQ-4 scores

Approximately 39% of the Sailors responded that their personal morale was high or very high, whereas 21% of the Sailors responded that their personal morale was low or very low. After assigning values to the responses, the average score was 2.21 ± 0.974, ranging from 0 to 4. In terms of unit morale, approximately 22% of the Sailors responded that their personal morale was high or very high, while 29% of the Sailors responded that their unit morale was low or very low. After assigning values to the responses, the average score was 1.87 ± 0.915, ranging from 0 to 4. Figures 17 and 18 show the distribution plots of personal and unit morale scores, respectively.

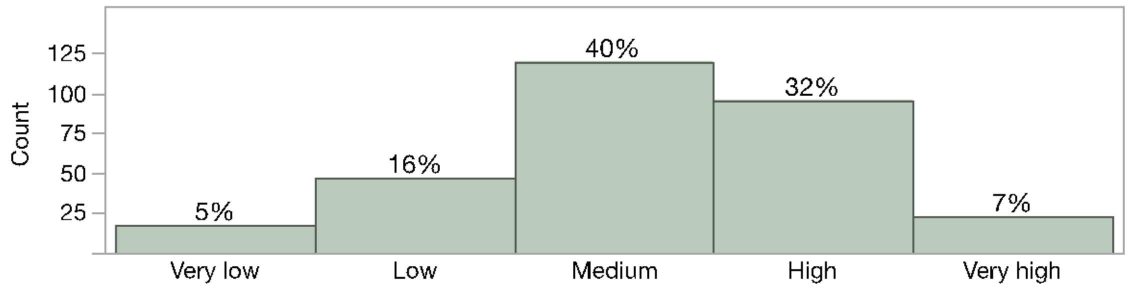


Figure 23. Distribution plot of responses in the NCCOSC item on personal morale

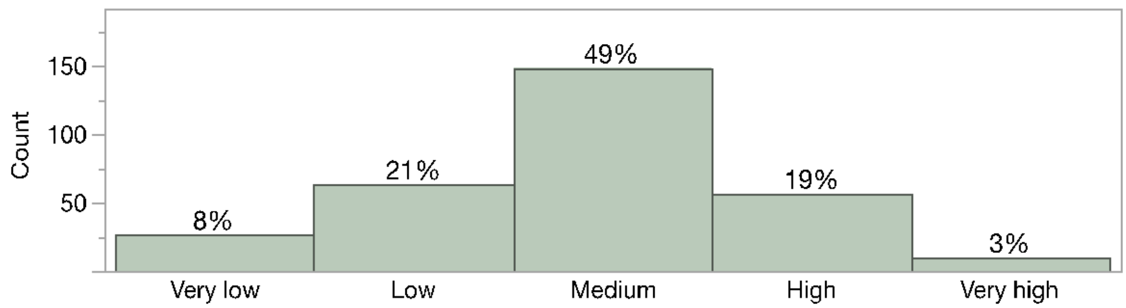


Figure 24. Distribution plot of responses in the NCCOSC item on unit morale

In terms of health status, approximately 39% of the Sailors responded that their health was in general very good or excellent, while 14% responded that their health was fair or poor. After assigning values to the responses, the average score was 2.32 ± 0.837, ranging from 0 to 4. Figure 19 shows the distribution plot of health status scores.

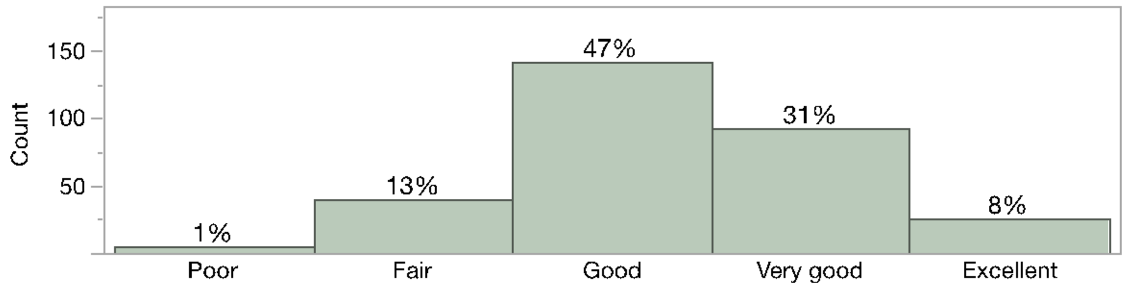


Figure 25. Distribution plot of SF-36 health responses

The following table shows an overview of the scores of the scales used in the SET. The table also includes the 25th, 50th, and 75th percentile scores.

Table 3. Scores of scales included in the SET

Measure	M ± SD	Range	Percentile 25th	Percentile 50th	Percentile 75th
Connor-Davidson Resilience Scale (CD-RISC)	28.1 ± 6.22	6 – 40	24	29	32
Response to Stressful Experiences Scale (RSES-4)	11.1 ± 2.82	1 – 16	11	12	14
Brief Cope (BC-18)					
Active coping	3.69 ± 1.47	0 – 6	3	4	5
Denial [1]	0.88 ± 1.35	0 – 5	0	0	1
Using emotional support	2.10 ± 1.65	0 – 6	1	2	3
Behavioral disengagement [1]	1.10 ± 1.34	0 – 6	0	1	2
Using instrumental support	2.51 ± 1.78	0 – 6	1	2	4
Positive reframing	3.56 ± 1.62	0 – 6	3	4	5
Self-blame [1]	2.17 ± 1.72	0 – 6	1	2	4
Planning	3.46 ± 1.65	0 – 6	2	4	5
Religion	1.86 ± 2.09	0 – 6	0	1	3
New General Self-Efficacy Scale (NGSES)	32.5 ± 5.06	8 – 40	30	32	36
Brief Self-Control Scale (BSS)	-12.0 ± 8.03	-33 – 11	-17	-12	-7
Life Orientation Test-Revised (LOT-R)	1.97 ± 4.33	-12 – 11	-1	2	5
Thought Control Q. (TCQ) – Reappraisal Subscale	9.86 ± 3.33	0 – 18	8	10	12
Cognitive Emotional Regulations Q. (CERQ) – Positive Reappraisal Subscale	14.1 ± 3.73	4 – 20	12	14	17
Unit Support Scale (USS)	40.2 ± 9.57	12 – 60	34	41.5	48
Morale					
Personal morale	2.20 ± 0.97	0 – 4	2	2	3
Unit morale	1.87 ± 0.92	0 – 4	1	2	2
The Satisfaction with Life Scale (SWLS)	22.2 ± 6.52	5 – 35	18	23	27
SF-36 – overall health condition	2.32 ± 0.84	0 – 4	2	2	3
Perceived Stress Scale (PSS) [1]	-1.17 ± 2.77	-8 – 7	-3	-1	0.75
Primary Care PTSD Screen (PC-PTSD) [1]	1.32 ± 1.46	0 – 4	0	1	2
Patient Health Questionnaire - 4 (PHQ-4) [1]	2.94 ± 2.78	0 – 12	1	2	4.75
Total Mood Disturbance (TMD) [1]	36.3 ± 34.5	-23 – 142	9	29.5	54
Tension-Anxiety[1]	9.27 ± 6.29	0 – 31	4	8	13
Depression[1]	11.2 ± 11.7	0 – 60	2	8	16
Anger-hostility[1]	11.7 ± 9.33	0 – 44	4	9	18
Vigor	13.3 ± 6.19	0 – 32	9	13	18
Fatigue[1]	9.54 ± 6.32	0 – 26	5	8	14
Confusion - bewilderment[1]	7.99 ± 4.87	0 – 23	4	7	11

[1] Inverse scoring. Lower score denotes more favorable attribute.

B. QUALITATIVE ASSESSMENT OF SET

1. Focus of SET battery – Assessment of traits, states, or both?

According to the supporting literature, most of the scales in the SET battery assess traits -- that is, personal characteristics that are enduring or stable across time. For example, the Cognitive Emotional Regulations Questionnaire (CERQ), the Life Orientation Test – Revised (LOT-R), the Response to Stressful Experiences Scale (RSES-4), and the Brief COPE (BC) questionnaires all assess personal traits. In contrast, POMS explicitly assesses states, that is, characteristics that are situation-dependent. Of note, four scales included instructions that led the participants to respond based on their experiences in the previous month (The Connor-Davidson Resilience Scale – CD-RISC, the Perceived Stress Scale – PSS, the Primary Care Posttraumatic Stress Disorder – PC-PTSD) or in the previous two weeks (the Patient Health Questionnaire – PHQ). The previous information suggests that, even though some questionnaires are state-oriented, the SET is predominantly a tool to assess indwelling traits.

2. SET components

The version of SET used in the study included 25 component scores derived from 17 scales, i.e., 11 validated questionnaires, 3 subscales from validated questionnaires, and 3 independent items.

We identified three issues of concern. First, the SET includes three independent items that focused on personal morale, unit morale, and one question which was extracted from the SF-36 and referred to condition of general health. From a psychometric perspective independent items are not recommended because they can provide unreliable results. In general, single-item scales should be examined psychometrically, and their psychometric quality and potential limitations should be appropriately assessed (Furr, 2011).

The second issue of concern is that the SET assesses facets of one's character which identify Sailors in the context of their profession. One's ability to cope with stress, personal morale, support from peers and unit leaders, and morale in the unit are sensitive issues. In such cases, questionnaire responses may be biased through the distorting lens of social desirability,i.e., the tendency of respondents to answer in a way considered positive and desirable (DeVellis, 2003; Edwards, 1957).

Our last concern was related to the decision to include specific overlapping scales in the SET. For example, the Connor-Davidson Resilience Scale (CD-RISC) and the Response to Stressful Experiences Scale - 4 items (RSES-4). According to the developers of the scale the CD-RISK (CD-RISC) assesses the ability to withstand, recover, grow and function competently in the face of stressors, adversity and changing demands (Campbell-Sills & Stein, 2007). The Response to Stressful Experiences Scale (RSES-4) assesses cognitive, emotional, and behavioral responses to stressful situations (De La Rosa et al., 2016). It is obvious that the two scales have overlap.Overlap also exists between the Profile of Mood States (POMS) and the Patient Health Questionnaire (Kroenke et al., 2009; McNair et al., 1971). The PHQ assesses psychological distress from depression and anxiety symptoms. The POMS, however, also includes the depression and the tension – anxiety subscales.

3. Observations from fielding the SET

The SET battery was eight pages long in its printed form. Given that it included 14 questionnaires and 3 independent items, Sailors had to read and understand 17 different instructions and respond to 166 questions, with each question having from 2 to 5 levels (that is, choosing 166 among 800 levels). Even though it was not possible to accurately assess the length of time needed to take the SET, our observations suggested that Sailors needed 15 minutes or more to complete the battery, with some of them needing up to 25 minutes. Often, Sailors did not have the time needed to complete the questionnaire uninterrupted. In such cases, Sailors were allowed to complete the SET in their free time and return the completed battery to the research team at their convenience. The fact that the SET was too long was also evident by the missing data (incomplete questionnaires or items) which became more frequent towards the end of the battery.

C. QUANTITATIVE ASSESSMENT OF SET

1. Construct validity

Table 3 shows the results of inter-correlation analysis among the SET components based on Pearson’s r. Correlations ranged from -0.62 to 0.73. In terms of magnitude, approximately 42% of the correlations were ≤ 0.20, 23 correlations ≥ 0.5, while only 7 correlations exceeded 0.6 with 3 correlations were higher than 0.7). These correlations are shown in the grey color in Table 2. In terms of magnitude, the highest correlations were between CD-RISK and RSES-4 (r = 0.73), PHQ and POMS TMD (r = 0.72), CD-RISK and NGSES (r = 0.70), BC-18 Using Emotional Support and BC-18 Using Instrumental Support (r = 0.69), NGSES and RSES-4 (r = 0.63), and CD-RISK and PSS-4 (r = -0.62). There are two issues of note. First, the association between the BC-18 subscales is not unexpected given that both focus on some aspect of using support. Second, the associations between CD-RISK and other components of SET are not surprising if we consider that CD-RISK assesses general resilience.

After adjusting for components that have inversed scoring, the average intercomponent correlation was 0.23. This finding suggests that on average, the SET has an acceptable convergent validity (Piedmont, 2014) in the sense that the SET components are reasonably homogenous in assessing facets of resilience.

The pattern of our findings also shows, however, that the SET has good divergent properties. First, the average intercomponent correlation was close to 0.20, which is on the low side of criterion needed to accept scales measuring the same construct (Piedmont, 2014). Assuming that the SET components measure different facets of resilience, the average intercomponent correlation should be, and is, on the low side. Second, the correlations between the different components of SET show a wide range, with most of the correlations being weak.

Table 4. Correlation among SET components

SET component	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
(1) POMS TMD [1]																								
(2) CD-RISC	**-0.38**																							
(3) RSES-4	**-0.26**	**0.73**																						
(4) BC-18 Active coping	-0.03	**0.34**	**0.39**																					
(5) BC-18 Denial [1]	**0.45**	**-0.27**	**-0.20**	0.07																				
(6) BC-18 Using emotional support	0.05	0.04	**0.10**	**0.36**	**0.24**																			
(7) BC-18 Behavioral disengagement [1]	**0.47**	**-0.40**	**-0.32**	-0.06	**0.51**	**0.20**																		
(8) BC-18 Using instrumental support	**0.10**	-0.02	0.04	**0.31**	**0.17**	**0.69**	**0.12**																	
(9) BC-18 Posit. Reframing	-0.05	**0.30**	**0.40**	**0.51**	**0.09**	**0.28**	-0.03	**0.32**																
(10) BC-18 Self-blame [1]	**0.49**	**-0.32**	**-0.25**	0.04	**0.35**	**0.16**	**0.38**	**0.22**	0.04															
(11) BC-18 Planning	**0.12**	**0.19**	**0.29**	**0.58**	**0.20**	**0.39**	**0.09**	**0.46**	**0.52**	**0.30**														
(12) BC-18 Religion	0.06	0.08	**0.16**	**0.22**	**0.13**	**0.20**	0.07	**0.14**	**0.18**	-0.04	**0.24**													
(13) NGSES score	**-0.28**	**0.70**	**0.63**	**0.38**	**-0.18**	0.05	**-0.29**	0.05	**0.30**	**-0.32**	**0.22**	**0.14**												
(14) BSC score	**-0.26**	**0.33**	**0.36**	**0.18**	**-0.14**	-0.07	**-0.16**	-0.08	0.10	**-0.32**	**0.13**	**0.25**	**0.40**											
(15) LOT-R-6	**-0.38**	**0.51**	**0.48**	**0.23**	**-0.19**	0.11	**-0.30**	**0.18**	**0.31**	**-0.30**	**0.19**	**0.19**	**0.53**	**0.33**										
(16) TCQ-Reappraisal	0.06	**0.25**	**0.30**	**0.24**	0.05	0.01	-0.03	0.04	**0.35**	0.07	**0.27**	0.10	**0.24**	0.06	**0.14**									
(17) CERQ – Positive Reappraisal Subscale	**-0.21**	**0.55**	**0.60**	**0.38**	**-0.14**	0.09	**-0.26**	**0.12**	**0.43**	**-0.18**	**0.30**	**0.16**	**0.52**	**0.31**	**0.51**	**0.40**								
(18) USS	**-0.40**	**0.40**	**0.30**	**0.19**	**-0.23**	**0.13**	**-0.32**	**0.14**	**0.25**	**-0.22**	**0.18**	-0.01	**0.31**	**0.15**	**0.40**	0.01	**0.26**							
(19) Personal morale	**-0.51**	**0.45**	**0.41**	**0.14**	**-0.19**	**0.18**	**-0.28**	**0.17**	**0.24**	**-0.26**	**0.10**	0.04	**0.38**	**0.21**	**0.46**	0.08	**0.32**	**0.48**						
(20) Unit morale	**-0.36**	**0.28**	**0.21**	0.05	**-0.24**	0.11	**-0.31**	**0.15**	0.09	**-0.17**	0.06	-0.01	**0.21**	**0.15**	**0.30**	0.05	0.11	**0.46**	**0.46**					
(21) SWLS	**-0.45**	**0.39**	**0.38**	**0.17**	**-0.23**	0.06	**-0.30**	0.04	**0.17**	**-0.28**	0.08	0.09	**0.38**	**0.29**	**0.45**	0.04	**0.38**	**0.43**	**0.51**	**0.34**				
(22) SF-36 General health	**-0.33**	**0.45**	**0.37**	**0.22**	**-0.22**	**0.12**	**-0.21**	**0.15**	**0.24**	**-0.15**	**0.18**	0.11	**0.36**	**0.25**	**0.41**	0.08	**0.35**	**0.34**	**0.43**	**0.24**	**0.35**			
(23) PSS-4 [1]	**0.57**	**-0.62**	**-0.49**	**-0.16**	**0.42**	0.01	**0.47**	0.01	**-0.18**	**0.38**	-0.08	-0.09	**-0.50**	**-0.31**	**-0.58**	-0.01	**-0.47**	**-0.46**	**-0.46**	**-0.32**	**-0.55**	**-0.43**		
(24) PC-PTSD [1]	**0.37**	**-0.24**	**-0.20**	-0.05	**0.26**	0.01	**0.31**	0.02	**-0.14**	**0.25**	0.07	-0.05	**-0.16**	**-0.22**	**-0.29**	0.06	**-0.20**	**-0.20**	**-0.17**	-0.09	**-0.28**	**-0.23**	**0.42**	
(25) PHQ-4 [1]	**0.72**	**-0.37**	**-0.27**	-0.07	**0.44**	0.07	**0.44**	0.08	-0.08	**0.47**	**0.14**	0.02	**-0.26**	**-0.22**	**-0.40**	**0.12**	**-0.23**	**-0.38**	**-0.47**	**-0.26**	**-0.44**	**-0.33**	**0.54**	**0.47**

Post-hoc statistical significance assessed with BH-FDR procedure

Correlations in bold have $p < 0.05$.

[1] SET component with inverse scoring.

a. Latent constructs

We used exploratory factor analysis (EFA) to assess any underlying latent constructs in the SET battery. Based on Cattell's scree test ,we identified the point at which the data curve flattened out. Given that two eigenvalues were clustered before the bend, we conducted the EFA both using a 4-factor and a 5-factor model. All five factors had an eigenvalue greater than 1 (Kaiser criterion) (Kaiser, 1960). The 5-factor model led to a latent factor with only two components, whereas two other factors had three components. In contrast, the 4-factor model had only one factor with three components and none had less than three components. Furthermore, the communalities did not change substantively between the 4- and the 5-factor model. For these reasons, the 4-factor model was chosen for further analysis. In terms of eigenvalues, the first factor had an eigenvalue of 7.42, the second 3.60, the third 18.3, and the fourth 1.25.

Analysis of factor intercorrelations showed all four latent factors were correlated at a statistically significant level. More pronounced were the correlations between Factor 1 and Factors 2/4, and between Factor 2 and Factor 4. Factor intercorrelations are shown in Table 5 and Figure 26.

Table 5. Intercorrelations among factors

Factors	(1)	(2)	(3)
(1) Factor 1: Disturbance			
(2) Factor 2: Self-efficacy and response to stress	-0.443		
(3) Factor 3: Active engagement	-0.163	0.170	
(4) Factor 4: Morale	-0.591	0.394	0.106

Post-hoc statistical significance assessed with BH-FDR procedure

The first factor included components associated with the perceived stress and disturbance associated with stress, i.e., the Patient Health Questionnaire – 4 (PHQ-4), the Profile of Mood States (POMS) Total Mood Disturbance (TMD), the Primary Care PTSD Screen (PC-PTSD), the Brief Cope (BC-18) Self-blame subscale, the Perceived Stress Scale (PSS), and the Brief Cope (BC-18) Denial subscale. Of note, all these components of SET had inverse scoring (a higher score on these scales denoted a less favorable attribute).

The second factor included components of self-efficacy and response to stress, i.e., the Connor-Davidson Resilience Scale (CD-RISC), the Response to Stressful Experiences Scale (RSES-4), the New General Self-Efficacy Scale (NGSES), and the Positive Reappraisal Subscale of the Cognitive Emotional Regulations Questionnaire (CERQ), and the Reappraisal subscale from the Thought Control Questionnaire (TCQ).

The third factor included components of active engagement, i.e., the BC-18 Planning subscale, the BC-18 Positive reframing subscale, the BC-18 Active coping subscale, the BC-18 Using emotional support subscale, and the BC-18 Using instrumental support subscale. Of note, the BC-18 Positive reframing and the BC-18 Active coping subscale had a loading of greater than 0.40 in the second factor. Conceptually, however, we believe that they are a better fit to the third factor. Lastly, the fourth factor included components of morale, i.e., the unit morale item, the Unit Support Scale (USS), and the personal morale item.

Communalities, i.e., the proportion of variance explained by the factors, ranged from 0.115 to 0.748. Specifically, the components with the lowest communalities were BC-18 Religion (0.115), BSS (0.222), TCQ-Reappraisal (0.238), PC-PTSD (0.315), SF-36 general health condition (0.319), and BC-18 Denial (0.345).

Six components did not have a factor loading of 0.40 or more, i.e., the BC-18 Behavioral disengagement, the Life Orientation Test-Revised (LOT-R), and the Brief Self-Control Scale (BSS), the Satisfaction with Life Scale (SWLS), the BC-18 Religion subscale, and the SF-36 – overall health condition item. These results are shown in Table 6 and Figure 26.

Table 6. EFA loadings for the four first order factors after promax oblique rotation

Measure	Factor 1	Factor 2	Factor 3	Factor 4	Communalities
Factor 1: Disturbance					
Patient Health Questionnaire - 4 (PHQ-4)	**0.943**	0.153	-0.061	0.018	0.748
POMS TMD	**0.755**	0.136	0.027	-0.190	0.690
Primary Care PTSD Screen (PC-PTSD)	**0.667**	0.018	-0.113	0.197	0.315
BC-18 Self-blame	**0.477**	-0.081	0.250	-0.038	0.386
Perceived Stress Scale (PSS)	**0.451**	-0.306	-0.032	-0.213	0.630
BC-18 Denial	**0.402**	-0.003	0.264	-0.140	0.345
Factor 2: Self-efficacy and response to stress					
Connor-Davidson Resilience Scale (CD-RISC)	-0.057	**0.722**	-0.104	0.217	0.729
Response to Stressful Experiences Scale (RSES-4)	-0.011	**0.775**	0.008	0.094	0.678
New General Self-Efficacy Scale (NGSES)	0.014	**0.715**	-0.046	0.163	0.607
Cognitive Emotional Regulations Q. (CERQ) – Positive Reappraisal Subscale of the	-0.101	**0.670**	0.128	-0.036	0.537
Thought Control Questionnaire (TCQ) -Reappraisal	0.275	**0.537**	0.025	-0.050	0.238
Factor 3: Active engagement					
BC-18 Planning	0.158	0.370	**0.587**	-0.046	0.551
BC-18 Positive reframing	-0.080	**0.412**	**0.480**	-0.113	0.444
BC-18 Active coping	-0.103	**0.449**	**0.518**	-0.202	0.505
BC-18 Using emotional support	0.034	-0.149	**0.747**	0.209	0.601
BC-18 Using instrumental support	0.087	-0.169	**0.764**	0.276	0.666
Factor 4: Morale					
Unit morale item	0.078	-0.051	-0.023	**0.721**	0.434
Unit Support Scale (USS)	-0.143	0.071	0.102	**0.506**	0.422
Personal morale item	-0.192	0.077	0.110	**0.557**	0.548
SET components with factor loading < 0.40					
BC-18 Behavioral disengagement	0.362	-0.154	0.219	-0.171	0.381
Life Orientation Test-Revised (LOT-R)	-0.250	0.367	0.136	0.223	0.489
Brief Self-Control Scale (BSS)	-0.212	0.368	-0.039	-0.057	0.222
The Satisfaction with Life Scale (SWLS)	-0.317	0.160	0.052	0.297	0.411
BC-18 Religion	-0.060	0.205	0.257	-0.162	0.115
SF-36 – overall health condition item	-0.199	0.241	0.137	0.232	0.319

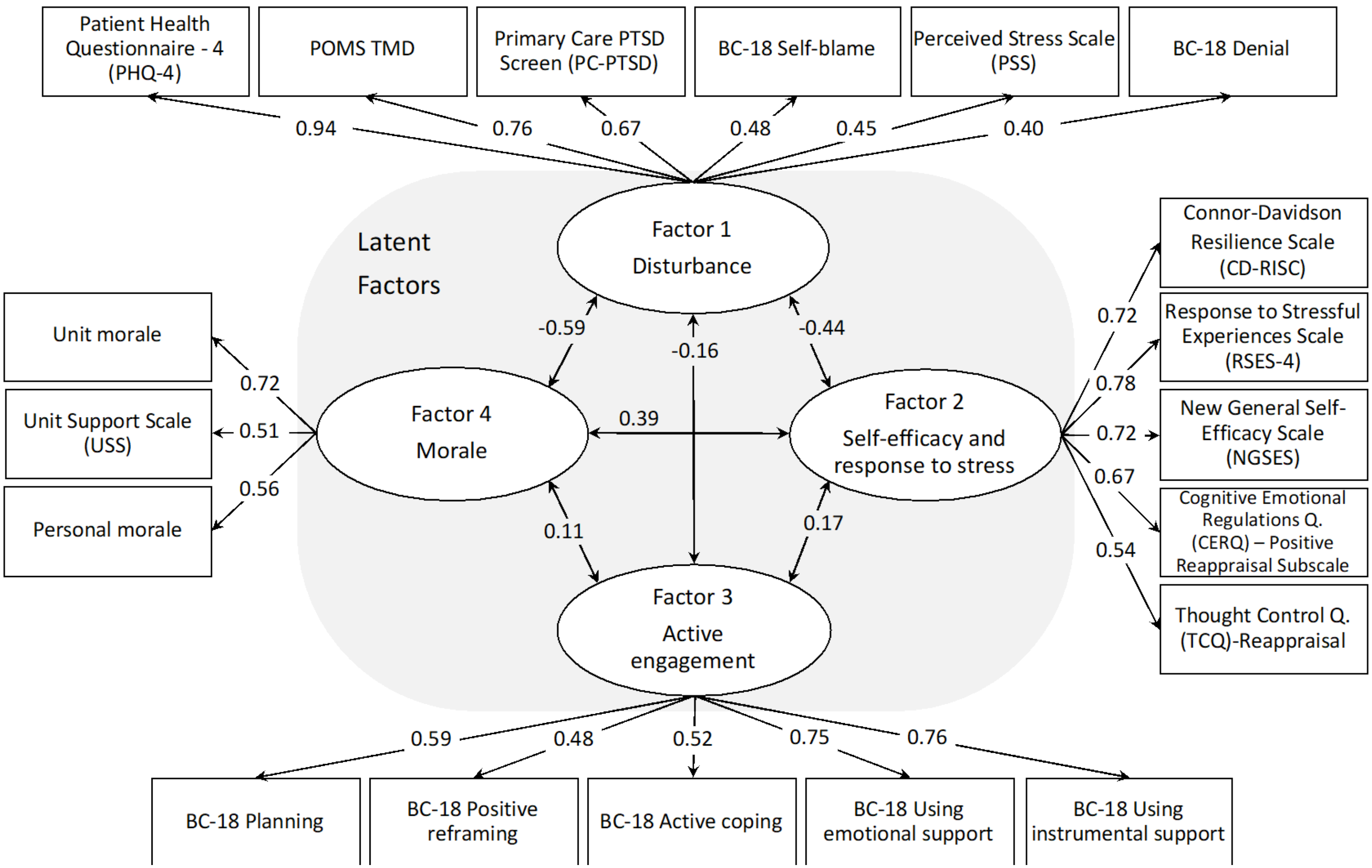


Figure 26. Results of the Exploratory Factor Analysis (EFA)

D. ASSOCIATION BETWEEN SET SCORES AND DEMOGRAPHIC AND OCCUPATIONAL VARIABLES

This section will address the criterion validity of the SET battery, i.e., the association between SET and variables of interest (DeVellis, 2003; Litwin, 2003).

1. Correlation between SET components and other variables

Based on Pearson's r, Table 6 shows pairwise correlations between SET components and other variables of interest. Results showed three interesting patterns. First, the average correlation between morningness-eveningness (ME) scores and the SET component scores was 0.13. Out of 25 SET component scores, 15 were correlated with ME scores at a statistically significant level ($p < 0.05$) with an average correlation of $r = 0.19$. These findings suggest that, in general, Sailors scoring higher in morningness had more favorable scores in terms of the SET components. In terms of magnitude, the highest correlations were between ME and BSS score ($r = 0.30$), CD-RISK score ($r = 0.25$), and NGSES score ($r = 0.25$).

Second, the average correlation between ISI scores and the SET component scores was -0.18. Out of 25 SET components, 15 were correlated with ME scores at a statistically significant level ($p < 0.05$) with an average correlation of $r = -0.29$. These findings suggest that, in general, Sailors with more severe insomnia symptoms had less favorable scores in terms of the corresponding SET components. In terms of magnitude, the highest correlations were between the ISI and POMS TMD ($r = 0.54$), and ISI and PHQ scores ($r = 0.43$). Lastly, the average correlation between PSQI scores and the SET component scores was -0.15. Out of 25 SET components, 15 were correlated with PSQI Global scores at a statistically significant level ($p < 0.05$) with an average correlation of $r = -0.23$. These findings suggest that, in general, Sailors with worse sleep quality had less favorable scores in terms of the corresponding SET components. In terms of magnitude, the highest correlations were between the ISI and POMS TMD ($r = 0.42$), and ISI and PHQ scores ($r = 0.35$). Overall, the pattern of pair-wise correlations shows that, in general, some SET components were associated with ME, ISI, and PSQI, but these associations are on average small.

Table 7. Correlations between SET components and other variables

Variable	Age (26)	ME (27)	ESS (28)	ISI (29)	PSQI (30)	Daily sleep duration (31)	Sleep episodes per day (32)
(1) POMS TMD [1]	-0.09	**-0.13**	**0.24**	**0.54**	**0.42**	-0.04	0.12
(2) CD-RISC	0.09	**0.25**	-0.08	**-0.23**	**-0.19**	-0.02	-0.10
(3) RSES-4	**0.12**	**0.18**	-0.05	-0.11	**-0.12**	-0.05	-0.04
(4) BC-18 Active coping	**0.12**	0.04	-0.01	-0.03	0.02	**-0.12**	-0.04
(5) BC-18 Denial [1]	-0.09	-0.09	0.06	0.30	0.13	-0.09	0.08
(6) BC-18 Using emotional support	-0.04	-0.01	0.05	0.03	-0.05	-0.09	0.02
(7) BC-18 Behavioral disengagement [1]	-0.10	-0.09	0.08	**0.29**	**0.20**	0.04	0.04
(8) BC-18 Using instrumental support	0.05	0.03	0.06	0.07	-0.03	-0.02	-0.07
(9) BC-18 Positive reframing	-0.06	0.10	0.05	-0.05	-0.06	-0.09	-0.08
(10) BC-18 Self-blame [1]	0.01	-0.04	0.06	**0.24**	**0.15**	0.02	0.01
(11) BC-18 Planning	0.09	0.08	-0.04	0.09	0.01	**-0.13**	-0.11
(12) BC-18 Religion	0.06	**0.16**	0.06	0.10	0.01	0.02	0.06
(13) NGSES	**0.12**	**0.25**	-0.07	**-0.14**	-0.07	-0.01	-0.03
(14) BSS	**0.13**	**0.30**	**-0.18**	**-0.13**	-0.06	-0.05	-0.08
(15) LOT-R-6	**0.17**	**0.21**	**-0.13**	**-0.23**	**-0.19**	-0.09	**-0.14**
(16) TCQ – Reappraisal	**-0.14**	-0.07	**0.13**	0.07	-0.02	0.06	-0.08
(17) CERQ – Positive Reappraisal	0.10	**0.15**	-0.03	-0.07	**-0.12**	-0.05	-0.07
(18) USS	0.08	**0.13**	-0.12	**-0.28**	**-0.22**	**-0.15**	**-0.16**
(19) Personal morale item	**0.12**	**0.13**	-0.08	**-0.30**	**-0.30**	0.05	-0.08
(20) Unit morale item	**0.17**	**0.14**	**-0.14**	**-0.37**	**-0.33**	-0.01	**-0.17**
(21) SWLS	**0.24**	**0.22**	-0.05	**-0.25**	**-0.17**	-0.10	-0.06
(22) SF-36 – overall health condition item	0.06	**0.21**	-0.09	**-0.23**	**-0.22**	-0.11	-0.03
(23) PSS-4 [1]	**-0.15**	**-0.23**	**0.17**	**0.35**	**0.28**	0.08	0.08
(24) PC-PTSD [1]	-0.06	**-0.15**	**0.17**	**0.29**	**0.23**	-0.07	0.07
(25) PHQ-4 [1]	**-0.14**	-0.11	**0.14**	**0.43**	**0.35**	-0.04	0.05
(26) Age		**0.17**	-0.10	-0.06	0.03	0.06	-0.05
(27) ME			-0.11	**-0.20**	-0.12	0.01	-0.05
(28) ESS				0.29	0.11	-0.10	0.04
(29) ISI					0.64	-0.06	0.25
(30) PSQI						-0.03	0.23
(31) Daily sleep duration							0.03

Post-hoc statistical significance assessed with BH-FDR procedure
Correlations in bold have $p < 0.05$.
[1] SET component with inverse scoring.

2. Associations between SET composite score(s) and other variables

So far, we have assessed the psychometric properties of SET based on the SET components. In this section of the report, we explore the utility of using one or more composite scores extracted from the SET components and use this/these score(s) to assess the criterion validity of SET. We should caution against drawing conclusions based on this analysis for reasons that will be explained later.

a. One composite score (unidimensional approach)

Initially, we developed a crude unidimensional composite score for the SET. First, the score of each SET component was normalized from 0 to 1. The composite score was calculated as the average of the normalized component scores. The components with inverse scoring (POMS TMD, BC-18 denial, BC-18 behavioral disengagement, BC-18 self-blame, PSS-4, PC-PTSD, and PHQ-4; higher scores denote less favorable attribute) were subtracted in the calculation of the overall SET score. Hence, the composite SET score ranged from 0 to 100 with 100 denoting higher resilience. As shown in Figure 27, the average composite score was 60.7 ± 11.2 ranging from 23.3 to 92.8.

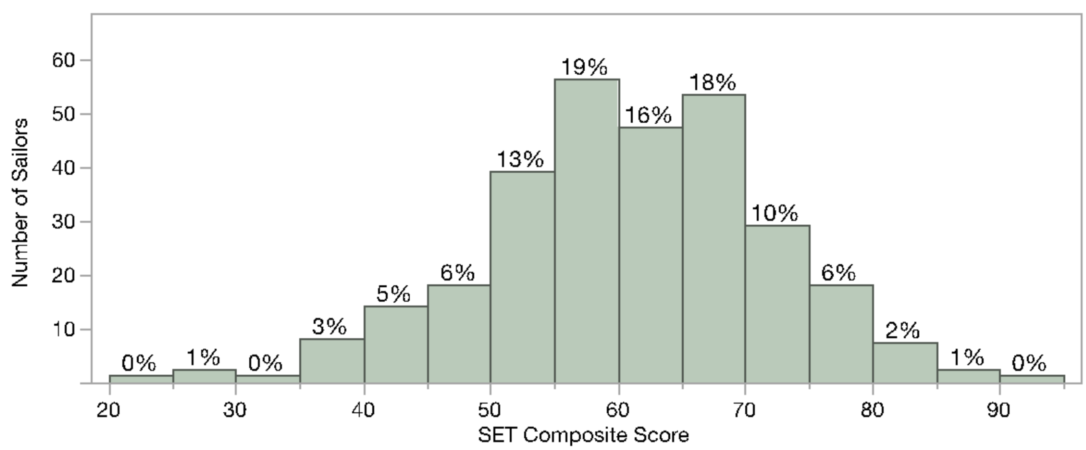


Figure 27. Distribution plot of SET Composite score

Multiple regression analysis was used to assess whether the SET composite score was associated with operationally relevant factors (e.g., characteristics of watchstanding schedules and sleep attributes). The first model focused on watchstanders working on 3- and 4- section watchbills, and included watchbill type (fixed, rotating) and number of sections in the watchbill as the potential predictor factors. Results showed that the model was not statistically significant ($F(6, 172) = 1.65$, $p = 0.137$; $R^2 = 0.054$).

The second model included all 296 Sailors with sleep attributes (daily sleep duration, and number of sleep episodes per day) as the potential predictor factors. Results showed that the unidimensional Composite score (model: $F(6, 254) = 3.04$, $p = 0.007$; $R^2 = 0.067$) was not associated with daily sleep duration ($p = 0.464$) or number of sleep episodes per day ($p = 0.088$). Officers had higher scores than Enlisted personnel ($p < 0.001$).

b. Composite scores after dimensionality reduction (multidimensional approach)

In order to reduce the number of component factors, we used Principal Component Analysis (PCA) with factors rotated with the promax method (oblique). The first factor was characterized by components associated with perceived stress, disturbance associated with stress, morale, support, and health. These included the Profile of Mood States (POMS) Total Mood Disturbance (TMD), the Patient Health Questionnaire (PHQ-4), the Perceived Stress Scale (PSS), the BC-18 Behavioral disengagement subscale, the Brief Cope (BC-18) Denial subscale, the Brief Cope (BC-18) Self-blame subscale, the SF-36 – overall health condition item, the Life Orientation Test-Revised (LOT-R), the Satisfaction with Life Scale (SWLS), the unit morale item, the Unit Support Scale (USS), and the personal morale item.

The second factor was characterized by components of active engagement, response to stress, and self-efficacy, i.e., the BC-18 Planning subscale, the BC-18 Positive reframing subscale, the BC-18 Active coping subscale, the Positive Reappraisal subscale from the Cognitive Emotional Regulations Questionnaire (CERQ), the Response to Stressful Experiences Scale (RSES-4), the New General Self-Efficacy Scale (NGSES), and the Connor-Davidson Resilience Scale (CD-RISC). The third factor was characterized by components of active engagement, and using support, i.e., the BC-18 Planning subscale, the BC-18 Active coping subscale, the BC-18 Using emotional support subscale, and the BC-18 Using instrumental support subscale. Lastly, the fourth factor included three components with loadings of more than 0.40, i.e., the Brief Self-Control Scale (BSS), the BC-18 Religion subscale, and the Primary Care PTSD Screen (PC-PTSD). Detailed results are shown in Table 8.

Table 8. PCA loading factors

SET component	Factor 1	Factor 2	Factor 3	Factor 4	Communalities
POMS TMD [1]	**0.811**	0.141	0.086	-0.082	0.672
PHQ-4 [1]	**0.762**	0.178	0.068	-0.196	0.658
PSS-4 [1]	**0.647**	-0.204	0.052	-0.170	0.668
BC-18 Behavioral disengagement [1]	**0.582**	-0.175	0.315	0.144	0.501
BC-18 Denial [1]	**0.571**	-0.014	0.365	0.145	0.468
BC-18 Self-blame [1]	**0.423**	0.047	0.338	-0.319	0.509
PC-PTSD [1]	0.394	0.153	0.008	**-0.411**	0.377
TCQ – Reappraisal	0.260	0.79	-0.116	-0.264	0.534
BC-18 Religion	0.251	0.044	0.286	0.**692**	0.522
BC-18 Planning	0.146	**0.462**	**0.575**	0.005	0.638
BC-18 Active coping	0.064	**0.499**	**0.412**	0.136	0.526
BC-18 Positive reframing	-0.008	**0.531**	0.392	0.005	0.516
BSS	-0.053	0.211	-0.097	**0.587**	0.495
CERQ – Positive Reappraisal	-0.111	**0.709**	-0.009	0.109	0.624
BC-18 Using emotional support	-0.124	-0.137	**0.861**	0.089	0.697
BC-18 Using instrumental support	-0.155	-0.090	**0.856**	-0.032	0.703
RSES-4	-0.196	**0.710**	-0.050	0.101	0.686
NGSES	-0.210	**0.634**	-0.074	0.157	0.623
CD-RISC	-0.366	**0.643**	-0.123	0.026	0.711
SF-36 – overall health condition item	**-0.443**	0.194	0.186	0.104	0.367
LOT-R-6	**-0.446**	0.301	0.143	0.231	0.545
SWLS	**-0.595**	0.094	0.107	0.127	0.468
Unit morale item	**-0.704**	-0.007	0.197	-0.344	0.482
USS	**-0.719**	0.083	0.243	-0.221	0.542
Personal morale item	**-0.730**	0.089	0.245	-0.105	0.576

[1] SET component with inverse scoring.

Next, we calculated the four SET composite scores based on the factors of the PCA. Multiple regression analysis was used to assess the association between SET composite scores, sleep attributes and watchbill attributes. Specifically, regression models included watchbill type (fixed, rotating), number of sections in the watchbill, daily sleep duration, and number of sleep episodes per day as the potential predictor factors. Rank group (Officer, Enlisted) and ship were included as confounding factors. The first group of models focused on watchstanders working on 3- and 4- section watchbills, and included watchbill type (fixed or rotating) and number of sections in the watchbill as the potential predictor factors. Results showed that the scores of Factor 1 (model: $F(6, 171) = 2.25$, $p = 0.041$; $R^2 = 0.073$) were not associated with watchbill type ($p = 0.650$) or the number of sections in the watchbill ($p = 0.452$). Officers, however, had lower scores on Factor 1 than Enlisted personnel ($p = 0.004$). The models for the rest of the factors were not statistically significant (Factor 2 model: $F(6, 171) = 2.04$, $p = 0.063$; $R^2 = 0.067$; Factor 3 model: $F(6, 171) = 1.08$, $p = 0.375$; $R^2 = 0.037$; Factor 4 model: $F(6, 171) = 1.31$, $p = 0.255$; $R^2 = 0.044$).

The second group of models included all 296 Sailors with sleep attributes (daily sleep duration, and number of sleep episodes per day) as the potential predictor factors. Results showed that the scores on Factor 1 (model: $F(6, 253) = 4.96$, $p < 0.001$; $R^2 = 0.105$) were not associated with daily sleep duration ($p = 0.946$) or the number of sleep episodes per day ($p = 0.066$). Officers, however, had lower scores on Factor 1 than Enlisted personnel ($p < 0.001$). The models for the rest of the factors were not statistically significant (Factor 2 model: $F(6, 253) = 1.75$, $p = 0.110$; $R^2 = 0.040$; Factor 3 model: $F(6, 253) = 1.08$, $p = 0.373$; $R^2 = 0.025$; Factor 4 model: $F(6, 253) = 0.790$, $p = 0.579$; $R^2 = 0.018$).

c. Set Composite Score(s) differences among ships

Next, we assessed differences between ships in terms of SET component scores (uni- and multi-dimensional). Using multiple regression, the first group of models included ships (A, B, C, D) as the potential predictor factor, whereas the ship groups (mixed, circadian, non-circadian) was the potential predictor factor in the second group of models. In both cases, rank group (Officer, Enlisted) was the confounding factor.

Detailed results are shown in Tables 9 and 10. Overall, ship group (in terms of the dominant type of watchbill) was not associated with the SET composite score at a statistically significant level. Officers, however, had better scores (denoting higher resilience) than Enlisted personnel.

Table 9. Association between SET composite score(s) and ship

Composite score	Model p-value	Model R^2	Officer or Enlisted	Ship
Uni-dimensional	0.002	0.056	< 0.001	0.273
Multi-dimensional				
PCA Factor 1	< 0.001	0.086	< 0.001	0.094
PCA Factor 2	0.025	0.038	0.006	0.423
PCA Factor 3	0.441	0.013	-	-
PCA Factor 4	0.697	0.008	-	-

Inclusion criteria: p-values for potential predictor factors are shown when model p-value < 0.05.

Table 10. Association between SET composite score(s) and ship group

Composite score	Model p-value	Model R^2	Officer or Enlisted	Ship
Uni-dimensional	< 0.001	0.056	< 0.001	0.143
Multi-dimensional				
PCA Factor 1	< 0.001	0.082	< 0.001	0.081
PCA Factor 2	0.013	0.036	0.005	0.314
PCA Factor 3	0.309	0.012	-	-
PCA Factor 4	0.539	0.007	-	-

Inclusion criteria: p-values for potential predictor factors are shown when model p-value < 0.05.

d. Differences between Sailors in the highest and lower quartiles

Lastly, we assessed differences between Sailors who were in the highest and the lowest quartiles in terms of their unidimensional SET composite score. The highest quartile group included Sailors who scored above the 75th percentile of scores, i.e., 68.4%. The lowest quartile group included Sailors who scored below the 25th percentile of scores, i.e., 54.3%. Sailors with scores between the 25th and the 75th percentile were excluded from this analysis.

Table 11. Differences between Sailors in the highest and lowest quartile groups

Variable	Lowest quartile group	Highest quartile group	p-value
Age	26.0 ± 5.97	29.2 ± 7.20	0.003 [1,5]
Sex (males), # (%)	52 (70.3%)	58 (78.4%)	0.347 [2]
ME tendency, M ± SD	51.2 ± 8.69	56.6 ± 8.66	< 0.001 [1,5]
Enlisted, # (%)	66 (89.2%)	51 (68.9%)	0.004 [2,5]
Use nicotine products, # (%)	26 (35.1%)	22 (30.1%)	0.599 [2]
Use caffeinated beverages, # (%)	59 (79.7%)	62 (84.9%)	0.518 [2]
Sections in the watchbill, # (%) [3]			0.012 [2,5]
2	9 (12.2%)	1 (1.35%)	-
3	8 (10.8%)	3 (4.05%)	-
4	30 (40.5%)	42 (56.8%)	-
Rotating watchbill, # (%) [3]	13 (27.7%)	23 (50.0%)	0.034 [2,5]
ESS score, M ± SD	10.0 ± 4.84	9.11 ± 4.95	0.028 [1,5]
ESS > 10, # (%) [4]	39 (52.7%)	27 (36.5%)	0.069 [2,5]
ISI score, M ± SD	13.3 ± 4.45	8.66 ± 4.45	< 0.001 [1,5]
ISI score ≥ 15, # (%)	32 (43.2%)	8 (10.8%)	< 0.001 [2,5]
PSQI Global score, M ± SD	9.65 ± 2.93	6.89 ± 3.14	< 0.001 [1,5]
Poor sleepers, # (%)	66 (89.2%)	47 (63.5%)	< 0.001 [2,5]
Daily sleep duration, M ± SD	6.58 ± 1.07	6.44 ± 0.97	0.455 [1]
Number of sleep episodes per day, M ± SD	1.50 ± 0.39	1.38 ± 0.35	0.064 [1,5]

[1] Analysis of Variance

[2] Fisher’s Exact test

[3] Atypical Watchbills excluded (n = 55 Sailors)

[4] An ESS score of 10 or more indicates elevated daytime sleepiness

[5] Statistically significant according to post-hoc BH-FDR procedure

These results suggest that Sailors of younger age, Enlisted, less morningness prefernce, and Sailors standing watch in watchbills with fewer sections have lower SET scores (indicative of lower resilience) compared to Sailors who are older, Officers, and more morningness preference or Sailors who stand watch in watchbills with more sections. Also, Sailors who report worse sleep quality, more daytime sleepiness, and more severe insomnia symptoms have lower SET scores compared to Sailors who have better sleep quality, are less sleepy, and report less severe insomnia symptoms.

Of note, the Sailors working in rotating watchbills are overrepresented in the highest quartile group, approximately 2 to 1 ratio compared to the lowest quartile group. Further analysis, however, showed that result is misleading because most of the Sailors in the subgroup were from the same ship and were working on a 4-section watchbill (which has an advantage compared to the 3-section watchbill).

3. Differences between study phases (n = 61 Sailors)

The analysis presented thus far was based on the data of 296 Sailors who participated in Phase 1 of the study, that is, at the beginning of the deployment. This section of the report, however, focuses on comparing SET scores between the two Phases of the study, at the beginning and towards the end of the deployment. This comparison is based on 61 Sailors from two ships. The first ship (n = 31) was using circadian watchbills for both phases of the study. The second ship (n = 31) used both circadian and non-circadian watchbills in the beginning of the deployment, but only circadian watchbills towards the end. Overall, ~41% of the Sailors worked on circadian watchbills in Phase 1, ~13% on non-circadian watchbills, and ~46% did not stand watches. In Phase 2, ~59% of the Sailors worked on circadian watchbills, and ~41% did not stand watches.

We used a mixed-effects model analysis to identify differences between study phases. Ship was included as a confounding factor, and study phase was the potential predictor factor. Results showed a trend of worse scores in terms of PHQ-4, POMS TMD, PC-PTSD, SWLS, and USS. The same trend was identified for the SET unidimensional composite score, and the score of Factor 1 (multi-dimensional). These results are shown in Tables 12 and 13.

Table 12. The association between study phase and SET component/composite scores.

SET component	Un-adjusted p-value	Trend
POMS TMD	0.028	Worse scores in Phase 2
PHQ-4	0.004 [1]	Worse scores in Phase 2
PSS-4	0.216	-
BC-18 Behavioral disengagement	0.062	-
BC-18 Denial	0.849	-
BC-18 Self-blame	0.884	-
PC-PTSD	0.091	Worse scores in Phase 2
TCQ – Reappraisal	0.580	-
BC-18 Religion	0.167	-
BC-18 Planning	0.675	-
BC-18 Active coping	0.819	-
BC-18 Positive reframing	0.193	-
BSS	0.647	-
CERQ – Positive Reappraisal	0.800	-
BC-18 Using emotional support	0.600	-
BC-18 Using instrumental support	0.643	-
RSES-4	0.167	-
NGSES	0.764	-
CD-RISC	0.981	-
SF-36 – overall health condition item	0.606	-
LOT-R-6	0.803	-
SWLS	0.091	Worse scores in Phase 2
Unit morale item	0.626	-
USS	0.081	Worse scores in Phase 2
Personal morale item	0.168	-

Note: A description of the trend is provided when the p-value of the Study Phase is <0.10.
[1] Statistically significant according to post-hoc BH-FDR procedure

Table 13. The association between study phase and SET composite scores.

SET component	p-value	Factor 2
Uni-dimensional	0.072 [1]	Worse scores in Phase 2
Multi-dimensional		
PCA Factor 1	0.040 [1]	Worse scores in Phase 2
PCA Factor 2	0.707	-
PCA Factor 3	0.922	-
PCA Factor 4	0.292	-

Note: A description of the trend is provided when the p-value of the Study Phase is <0.10.
[1] Statistically significant according to post-hoc BH-FDR procedure

IV. DISCUSSION

Initially, the main purpose of this study was to contribute to the development of the Servicemember Evaluation Tool (SET) by assessing the criterion validity of the SET battery. Using data from a sample of Sailors performing their duties on USN ships, the focus was on the relationship between individual resilience -- as measured by the SET -- and occupational factors (e.g., sleep, work/rest schedules, destructive behaviors, etc.) The data we collected, however, combined with our observations from fielding the SET and a critical appraisal of the tool led the research team to extend the scope of this study.

First, we assessed the utility of the SET for collecting data in the field. Our observation was that the SET is extremely long for field studies with Sailors taking 15 minutes or more to complete the SET battery. The importance of the length of time needed to complete the SET is better understood if one considers the entire "recruitment – complete the questionnaire" process. Sailors are first recruited to the participate in the study, briefed on the specifics of the study, and their questions are answered. Next, Sailors read/sign the consent form, and then complete the questionnaires that include demographics and the SET *per se*. This entire process takes approximately 40 to 45 minutes to complete.

We also performed a critical appraisal of the SET battery in terms of its design (DeVellis, 2003; Spector, 1992). We identified a number of components which overlap. For example, the Connor-Davidson Resilience Scale (CD-RISC) and the Response to Stressful Experiences Scale (RSES), or the Profile of Mood States (POMS) and the Patient Health Questionnaire (PHQ). Given the length of the SET battery, it is important to reduce overlap by omitting scales which strongly correlate with each other.

The main issue of concern, however, was the conceptual construct and framework on which the SET is based and how well this was expressed in the components included in the SET battery (DeVellis, 2003; Ledesma, 2014; Spector, 1992). As already noted, the version of SET used in our study included 20 components which covered various individual resilience domains (i.e., optimism, positive coping, behavioral control, flexible thinking, control and confidence, effective coping with stressful situations;,life satisfaction, social support, emotional affect, perceived stress and psychological distress, information provided by NCCOSC). The 20 components, however, were only some of the components/scales that the developers had initially included in the SET. Therefore, SET is not a fixed battery with an internal logic based on a clear statement of what it truly measures. Of note, even though the SET battery is focused on assessing psychological resilience, this constitutes an overarching scope, not an explicit definition of what the SET measures. For example, it is not clearly defined whether SET assesses traits of resilience or states, or whether resilience is viewed as a process or an outcome (Southwick, Bonanno, Masten, Panter-Brick, & Yehuhda, 2014). The design of the SET battery would benefit also from distinguishing between external and internal variables or resilience, or on whether it is focused explicitly on individual resilience (Ledesma, 2014). These clarifications in terms of the framework on which the tool is based will then make it possible to identify the right scales to include, and whether it is appropriate to use a SET score or scores focused on individual resilience to assess unit/ship resilience.

Given these issues, we opted to omit using confirmatory analytical methods. Instead, we used an exploratory approach to assess the latent variables of the SET battery, its dimensionality, and some aspects of criterion validity to include developing composite scores. Therefore, our results regarding the association of SET components with demographic and other variables of interest should be interpreted as evidence that SET should be improved.

Based on this exploratory perspective, it is appropriate to discuss some of our findings. Overall, results show that the SET composite scores distinguish between Officers and Enlisted personnel with Officers scoring better (higher resilience) than Enlisted personnel. Given what is already known about the differences between circadian and non-circadian watchbills in terms of daytime sleepiness, severity of insomnia symptoms and psychomotor vigilance performance, it was expected that SET scores would differ between watchbill types. This hypothesis was not verified. That is, SET composite scores did not differ between watchbill types, by number of sections in the watchbills, and by sleep attributes. The pattern of results suggests that the SET is not sensitive to occupational factors which affect human states in terms of fatigue and sleepiness. If the SET batterys is focused on trait aspects of resilience then being insensitive to state factors is an expected behavior. Unfortunately, given the previous discussion regarding the focus of the SET battery, we cannot assess whether this (non) finding is a limitation of SET or a positive result.

Of course, it should be noted that the pattern of results when focusing on the Sailors with low SET scores (lowest 25%) compared to Sailors with high SET scores (highest 25%) is promising. Our results suggest that Sailors of younger age, Enlisted, less morning types, and Sailors standing watch in watchbills with fewer sections have lower SET scores (indicative of lower resilience) compared to Sailors who are older, Officers, and more morning types or Sailors who stand watch in watchbills with more sections. Also, Sailors who report worse sleep quality, more daytime sleepiness, and more severe insomnia symptoms have lower SET scores compared to Sailors who have better sleep quality, are less sleepy, and report less severe insomnia symptoms. The factthat these trends become evident only after we exclude Sailors scoring between the 25% and the 75% percentile emphasizes the need to improve the focus of the SET fbattery.

In conclusion, the SET battery is a much-needed instrument for assessing psychological resilience in active duty service members. The instrument in its current form has some positive attributes but it should be revised in order to refine its focus, improve its psychometric properties, and reduce potential sources of bias. Once the revision occurs, a new data collection effort on multiple naval vessels could be used to reassess its utility.

A. RECOMMENDATIONS

We clustered our recommendations in two groups. The first group ("Primary") includes the recommendations which we consider most important to improve the SET battery. The other groups ("Secondary") includes recommendations which focus more on the current version of the SET battery, and may not be appropriate if the SET is revised and its current scales change.

1. Primary recommendations

- Refine the theoretical construct and framework on ehich the SET battery is built, and reassess which scales and components the SET battery will include.

- Reduce the size of the SET battery.
- Use best practices regarding questionnaire design.

2. Secondary recommendations

- Integrate a method to account for social desirability (Ballard, 1992; Crowne & Marlowe, 1960; Fischer & Fick, 1993; Nederhof, 1985) and other potential sources of bias.
- Replace the three independent items (personal morale, unit morale, and the SF-36 question on general health) with validated instruments, or assess the psychometric properties of these items.
- Omit the Response to Stressful Experiences Scale (RSES-4), which overlaps (r = 0.73) with the Connor-Davidson Resilience Scale (CD-RISC) (Campbell-Sills & Stein, 2007; De La Rosa et al., 2016).
- Consider replacing the POMS with a validated but shorter instrument.
- If the POMS is included in the SET battery, consider omitting the Patient Health Questionnaire (Kroenke et al., 2009; McNair et al., 1971) which is highly correlated with POMS TMD (r = 0.72).
- Consider omitting the Perceived Stress Scale (PSS) because it is correlated (r > 0.50) with five other SET components, i.e., POMS (r = 0.57), CD-RISK (r = -0.62), NGSES (r = -0.50), SWLS (r = -0.55), and PHQ (r = 0.54).
- Consider omitting the New General Self-Efficacy Scale (NGSES) because it is correlated (r > 0.50) with five other SET components, i.e., CD-RISK (r = 0.70), RSES (r = 0.63), LOT-R (r = 0.53), CERQ (r = 0.52), and PSS (r = -0.50).

B. STUDY LIMITATIONS

This study had a number of limitations. The data we collected were not balanced between ships, that is, not all factors are equally represented in all ships.. For example, approximately 65% of Sailors working on a rotating watchbills and ~80% of the watchstanders in the rotating 4-section watchbills were from one ship, whereas the majority of watchstanders in the rotating 3-section watchbills are from a specific department (~80% Operations). This means that our SET results, in terms of the effect of watchbill, may be biased by the command climate on a specific ship. Future efforts should collect data from a larger sample of ships. This expansion will also help account for possible differences in the command climate between vessels.

C. ACKNOWLEDGEMENTS

This study could not have been accomplished without the support of the leadership and crews of the ships who participated in our study. We would like to thank CAPT Charles Good USN, Surface Warfare Chair, NPS for his assistance in organizing the moving pieces in this study, and the following individuals for assisting in collecting data on the ships:

- Operations Research Department, Naval Postgraduate School: Dr. Brennan Cox, LCDR, USN; Dr. Anna Sjörs Dahlman; Dr. Carl Magnus Dahlman; Christine Fletcher, LCDR, USN; Alexandra DeAngelis, CPT, US Army; Joseph Felix, LCDR, USN; Anthony Baldessari, ENS, USN; Joshua Strubel.
- Graduate School of Business and Public Policy, Naval Postgraduate School: Dr. Chad Seagren
- Naval Health Research Center: Dr. Uade Olaghere da Silva, LT, USN
- COMNAVAIRSYSCOM PAX: Dr. Aditya Prasad, LT, USN
- NAMRU Dayton, OH: Dr. Adam Biggs, LT, USN; Eric Vorm, LCDR, USN

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Adam, K., & Oswald, I. (1984). Sleep helps healing. *British Medical Journal, 289*, 1400-1401.

American Psychological Association. (2018). The road to resilience. Retrieved October 23, 2018, from https://www.apa.org/helpcenter/road-resilience.aspx

Ancoli-Israel, S., Martin, J. L., Blackwell, T., Buenaver, L., Liu, L., Meltzer, L. J., . . . Taylor, D. J. (2015). The SBSM Guide to Actigraphy Monitoring: Clinical and Research Applications. *Behavioral Sleep Medicine*, *13*(1), S4-S38.

Ballard, R. (1992). Short forms of the Marlowe-Crowne social desirability scale. *Psychological Reports*, *71*(3 Pt 2), 1155-1160.

Bastien, C. H., Vallieres, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Medicine, 2*(4), 297-307.

Bates, M. J., Fallesen, J. J., Huey, W. S., Packard, G. A. J., Ryan, D. M., Burke, C. S., . . . Bowles, S. V. (2013). Total Force Fitness in units part 1: Military demand-resource model. *Military Medicine, 178*(11), 1164-1182. doi: 10.7205/MILMED-D-12-00519

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*, 289–300.

Brown, S., Matsangas, P., & Shattuck, N. L. (2015a, October 26 - 30). *Comparison of a circadian-based and a forward rotating watch schedules on sleep, mood, and psychomotor vigilance performance* Paper presented at the Human Factors and Ergonomics Society (HFES) Annual Meeting, Los Angeles, CA.

Brown, S., Matsangas, P., & Shattuck, N. L. (2015b). Is a sailor's life for you? Aches and pains of U.S. Navy sailors. *Sleep*, *38*(Abstract Supplement), A89.

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.

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.

Casey, G. W. J. (2011). Comprehensive soldier fitness: A vision for psychological resilience in the U.S. Army. *American Psychologist*, *66*(1), 1-3.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245-276.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational research methods*, *4*(1), 62-83.

Christian, L. M., Graham, J. E., Padgett, D. A., Glaser, R., & Kiecolt-Glaser, J. K. (2006). Stress an wound healing. *Neuroimmunomodulation*, *13*(5-6), 337-346. doi: 10.1159/000104862

Cohen, S., & Williamson, G. (1988). Perceived stress in a probability sample of the U.S. In S. Spacapam & S. Oskamp (Eds.), *The social psychology of health: Claremont Symposium on Applied Social Psychology* (pp. 31-67). Newbury Park, CA: Sage.

Cooper, C. L., Katona, C., & Livingston, G. (2008). Validity and reliability of the brief COPE in carers of people with dementia: The LASER-AD Study. *The Journal of Nervous and Mental Disease*, *196*(11), 838-843.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for fetting the most from your analysis. *Practical Assessment, Research and Evaluation*, *10*(7).

Crowne, D. P., & Marlowe, D. H. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349-354.

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*, *18*(3), 202-208.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA, US: SAGE Publications, Inc.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York, NY: Dryden Press.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272-299.

Fischer, D. G., & Fick, C. (1993). Measuring social desirability: Short forms of the Marlowe-Crowne social desirability scale. *Educational and Psychological Measurement*, *53*(2), 417-424.

Furr, R. M. (2011). Core principles, best practices, and an overview of scale construction *Scale Construction and Psychometrics for Social and Personality Psychology* (1st ed.). Thousand Oaks, CA, US: SAGE Publications, Ltd.

Garnefksi, N., & Kraaij, V. (2007). The cognitive emotion regulation questionnaire. *European Journal of Psychological Assessment*, *23*(3), 141-149.

Guo, S., & DiPietro, L. A. (2010). Factors affecting wound healing. *Journal of Dental Research*, *89*(3), 219-229. doi: 10.1177/0022034509359125

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*(3), 393-416.

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, *4*, 97–110.

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

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

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141-151.

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.

Lautenbacher, S., Kundermann, B., & Krieg, J.-C. (2006). Sleep deprivation and pain perception. *Sleep Medicine Reviews*, *10*(5), 357-369.

Ledesma, J. (2014). Conceptual frameworks and research models on resilience in leadership. *SAGE Open*, *4*(3). doi: 10.1177/2158244014545464

Litwin, M. S. (2003). *How to assess and interpret survey psychometrics* (2nd ed.): Sage Publications.
McNair, D. M., Lorr, M., & Droppelman, L. F. (1971). *Manual of the profile of mood states*. San Diego, CA: Educational and Industrial Testing Service.
Meltzer, L. J., Walsh, C. M., Traylor, J., & Westin, A. M. L. (2012). Direct comparison of two new actigraphs and polysomnography in children and adolescents. *Sleep, 35*, 159–166.
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.
Miller, N. L., Matsangas, P., & Kenney, A. (2012). The Role of Sleep in the Military: Implications for Training and Operational Effectiveness. In J. H. Laurence & M. D. Matthews (Eds.), *The Oxford Handbook of Military Psychology* (pp. 262–281). New York: Oxford University Press.
Miller, N. L., Matsangas, P., & Shattuck, L. G. (2008). Fatigue and its effect on performance in military environments. In P. A. Hancock & J. L. Szalma (Eds.), *Performance under stress* (1st ed., pp. 231–250). Burlington, VT: Ashgate Publishing.
Morgenthaler, T., Alessi, C., Friedman, L., Owens, J., Kapur, V. K., Boehlecke, B., . . . Swick, T. J. (2007). Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: An update for 2007. *Sleep*, *30*(4), 519-529.
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*, 601-608.
Mysliwiec, V., Matsangas, P., Baxter, T., & Shattuck, N. L. (2015). An Unusual Circadian Rhythm in an Active Duty Service Member. *Sleep and Biological Rhythms*, *14*(1), 113-115.
Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology, 15*(3), 263-280. doi: 10.1002/ejsp.2420150303
Nyenhuis, D. L., Yamamoto, C., Luchetta, T., Terrien, A., & Parmentier, A. (1999). Adult and geriatric normative data and validation of the Profile of Mood States. *Journal of Clinical Psychology, 55*, 79–86.
OPNAV. (2015). Navy total force manpower policies and procedures (OPNAVINST 1000.16L). Washington, D.C.: Department of the Navy.
Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological assessment*, *5*(2), 164.
Pedersen, E. R., Troxel, W. M., Shih, R. A., Pinder, E., Lee, D., & Geyer, L. (2015). Increasing resilience through promotion of healthy sleep among service members. *Military Medicine*, *180*(1), 4-6.
Peterson, C., Park, N., & Castro, C. A. (2011). Assessment for the U.S. Army Comprehensive Soldier Fitness program: The Global Assessment Tool. *American Psychologist*, *66*(1), 10-18. doi: 10.1037/a0021658
Piedmont, R. L. (2014). Inter-item correlations. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research*. Dordrecht: Springer.

Prins, A., O'uimette, 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), 9-14.

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*(2), 335-350.

Rupp, T. L., & Balkin, T. J. (2011). Comparison of Motionlogger Watch and Actiwatch actigraphs to polysomnography for sleep/wake estimation in healthy young adults. *Behavior Research Methods*, *43*, 1152-1160.

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

Shattuck, N. L., & Matsangas, P. (2014). *Work and rest patterns and psychomotor vigilance performance of crewmembers of the USS Jason Dunham: A comparison of the 3/9 and 6/6 watchstanding schedules* (Technical Report Report No. NPS-OR-14-004). Monterey, CA: Naval Postgraduate School.

Shattuck, N. L., & Matsangas, P. (2015a). A comparison of sleep and performance of U.S. Navy sailors on four different shiftwork schedules. *Sleep*, *38*(Abstract Supplement), A130.

Shattuck, N. L., & Matsangas, P. (2015b). Operational assessment of the 5-h on/10-h off watchstanding schedule on a US Navy ship: Sleep patterns, mood, and psychomotor vigilance performance of crew members in the nuclear reactor department. *Ergonomics*, *59*(5), 657-664. doi: 10.1080/00140139.2015.1073794

Shattuck, N. L., Matsangas, P., & Brown, S. (2015). *A comparison between the 3/9 and the 5/10 watchbills* (Technical Report Report No. NPS-OR-15-006). Monterey, CA: Naval Postgraduate School.

Shattuck, N. L., Matsangas, P., & Dahlman, A. S. (2018). Sleep and fatigue issues in military operations. In E. Vermetten, A. Germain, & T. Neylan (Eds.), *Sleep and Combat related PTSD*: Springer.

Shattuck, N. L., Matsangas, P., Moore, J., & Wegemann, L. (2015). *Prevalence of musculoskeletal symptoms, excessive daytime sleepiness, and fatigue in the crewmembers of a U.S. Navy ship* (Technical Report No. NPS-OR-15-005). Monterey, CA: Naval Postgraduate School.

Shattuck, N. L., Matsangas, P., Moore, J., & Wegemann, L. (2016). Prevalence of musculoskeletal symptoms, excessive daytime sleepiness and fatigue in the crewmembers of a U.S. Navy ship. *Military Medicine*, *181*(7), 655-662.

Shattuck, N. L., Matsangas, P., Mysliwiec, V., & Creamer, J. L. (2019). The role of sleep in human performance and well-being. In M. D. Matthews & D. Schnyer (Eds.), *Human Performance Optimization: The Science and Ethics of Enhancing Human Capabilities* (pp. 200-233). New York: Oxford University Press.

Shattuck, N. L., Matsangas, P., & Powley, E. H. (2015). *Sleep patterns, mood, psychomotor vigilance performance, and command resilience of watchstanders on the "five and dime" watchbill* (Technical Report No. NPS-OR-15-003). Monterey, CA: Naval Postgraduate School.

Southwick, S. M., Bonanno, G. A., Masten, A. S., Panter-Brick, C., & Yehuhda, R. (2014). Resilience definitions, theory, and challenges: Interdisciplinary

perspectives. *European Journal of Psychotraumatology, 5*(1). doi: 10.3402/ejpt.v5.25338
Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Thousand Oaks, CA, US: SAGE Publications, Inc.
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), 271-324.
Terman, M., Rifkin, J. B., Jacobs, J., & White, T. M. (2001). *Morningness-Eveningness Questionnaire (Revised)*. New York, NY: State Psychiatric Institute.
Troxel, W. M., Shih, R. A., Pedersen, E., Geyer, L., Fisher, M. P., Griffin, B. A., . . . Steinberg, P. S. (2015). *Sleep in the military: Promoting healthy sleep among U.S. servicemembers*. Santa Monica, CA: RAND.
Vogt, D. S., 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*(6), 710-717.
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.
Windle, G., Bennett, K. M., & Noyes, J. (2011). A methodological review of resilience measurement scales. *Health and Quality of Life Outcomes*, *9*(8). doi: 10.1186/1477-7525-9-8

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

3. Research Sponsored Programs Office, Code 41
   Naval Postgraduate School
   Monterey, CA 93943

4. Nita Lewis Shattuck
   Operations Research Department
   Naval Postgraduate School
   Monterey, California

5. Panagiotis Matsangas
   Operations Research Department
   Naval Postgraduate School
   Monterey, California

THIS PAGE INTENTIONALLY LEFT BLANK