



Afloat Surface Climate Assessment Survey Year One (FY21) Insights

Report prepared by:

Dale W. Russell, PhD
Commander, Naval Surface Forces
Commander, Naval Surface Force, U.S. Pacific Fleet (N05)
2841 Rendova Road
Coronado, CA 92155
dale.w.russell1.civ@us.navy.mil

DISTRIBUTION STATEMENT A. Approved for public release: distribution is unlimited.

**A – ADDED; M – MODIFIED; D – DELETED**

Document Version	Date	Section Modified	Change Type	Brief Description
1.0	20 Jan 2022	Entire Document	A	First version of the summary report, documenting results from Year One (FY21) of the ASCAS.



TABLE OF CONTENTS

TABLE OF CONTENTS	II
LIST OF TABLES	V
LIST OF FIGURES.....	VII
1 INTRODUCTION	9
1.1 Executive Summary and Year One Lessons Learned.....	9
2 BACKGROUND	12
2.1 Introduction.....	12
2.2 What is Safety Climate?.....	12
2.3 What is Crew Endurance?.....	13
3 SURVEY INSTRUMENT.....	14
3.1 Overview.....	14
3.2 Survey Measures.....	14
3.2.1 Demographics.....	14
3.2.2 Organizations Climate Perception Measures.....	14
3.2.3 Safety-Related Measures	16
3.2.4 Crew Endurance Measures	18
3.2.5 Leader Perceptions of Subordinates	20
3.2.6 Open-Ended Comments.....	20
4 RESULTS	21
4.1 Introduction.....	21
4.2 COMNAVSURFOR (Overall).....	21
4.2.1 Demographics.....	21
4.2.2 Areas of Praise and Concern	24
4.2.3 Responses to Key Constructs	24
4.2.4 Activity Analysis	26
4.3 CNSP/CNSL Analysis	27
4.3.1 Areas of Praise and Concern	27
4.3.2 Constructs with Notable Differences in Means.....	28
4.3.3 Responses to Key Constructs	28
4.3.4 Activity Analysis	29
4.4 Ship Class Analysis	30
4.4.1 Areas of Praise and Concern	30
4.4.2 Construct Deep Dive	31
4.4.2.1 Sleep Disturbance.....	31
4.4.2.2 Sleep Factors	32
4.4.2.3 Sleep Impairment.....	32
4.4.2.4 Job Stress	33



4.4.2.5	Job Demands	33
4.4.2.6	General Health.....	34
4.4.2.7	Mental Health	34
4.4.2.8	Physical Health.....	35
4.4.3	Breakdown of Sleep Factors.....	35
4.4.3.1	Sleep Factors of Highest Concern	36
4.4.3.2	Means Constructs Relating to Sleep.....	36
4.5	Paygrade Analysis.....	37
4.5.1	Areas of Praise And Concern	37
4.5.2	Constructs with Notable Differences in Means.....	37
4.5.3	Construct Deep Dive	38
4.5.3.1	Sleep Disturbance.....	39
4.5.3.2	Sleep Factors	39
4.5.3.3	Sleep Impairment.....	40
4.5.3.4	Job Stress	40
4.5.3.5	Job Demand.....	41
4.5.3.6	General Health.....	41
4.5.3.7	Mental Health	42
4.5.3.8	Physical Health.....	42
4.5.4	Breakdown of Sleep Factors.....	43
4.5.4.1	Sleep Factors of Highest Concern	43
4.5.4.2	Sleep Constructs of Highest Concern.....	43
4.5.5	Activity Analysis	44
4.6	Department Analysis.....	44
4.6.1	Areas of Praise by Department.....	45
4.6.2	Constructs with Noticeable Differences in Means	45
4.6.3	Construct Deep Dive	46
4.6.4	Breakdown of Sleep Factors.....	46
4.6.4.1	Sleep Factors of Highest Concern	46
4.6.4.2	Sleep Constructs of Highest Concern.....	47
4.6.5	Activity Analysis	47
4.7	Fit/Fill Analysis	48
4.7.1	Summary.....	48
4.7.2	Exploratory Analysis	48
4.7.2.1	Job Stress	50
4.7.2.2	Job Demands	51
4.7.2.3	Sleep Obtained.....	53
4.7.2.4	Sleep Deficiency.....	54
4.7.2.5	Workload as a Factor Affecting Sleep.....	55



4.7.2.6	Impairment Due to Fatigue.....	56
4.7.2.7	Sleep Disturbance.....	57
4.7.2.8	Refreshing Sleep.....	58
4.7.2.9	Time Available to Exercise.....	59
4.7.2.10	Leadership Assessment.....	60
4.7.2.11	Affective Commitment.....	61
4.7.2.12	Command Satisfaction.....	62
4.7.2.13	Command Safety Practices.....	63
4.7.2.14	Social Support.....	64
4.8	OFRP Analysis.....	65
4.9	Natural Language Processing (NLP) Analysis on Open-Ended Comments.....	65
4.9.1	Sentiment Analysis.....	65
4.9.2	Cluster Analysis.....	67
5	CONCLUSION.....	72
5.1	Summary.....	72
5.2	Next Steps.....	73
6	APPENDICES:.....	74
6.1	Appendix A: Ship Class Additional Analyses.....	74
6.2	Appendix B: Paygrade Mean Differences.....	78



LIST OF TABLES

<i>Table 1: Demographics of the ASCAS Respondents (SURFOR)</i>	22
<i>Table 2: Areas of Praise and Concern (SURFOR)</i>	24
<i>Table 3: Means of Constructs by Gender</i>	25
<i>Table 4: Means of Constructs by Paygrade</i>	25
<i>Table 5: Means of Constructs by Age</i>	26
<i>Table 6: CNSP/L Areas of Praise and Concern</i>	27
<i>Table 7: Constructs with Notable Differences in Means</i>	28
<i>Table 8: Means of Constructs based on Fleet Status and Tests of Significance between the Two Means</i>	29
<i>Table 9: Number of Ships and Sailors by Ship Class</i>	30
<i>Table 10: Ship Class Areas of Praise and Concern</i>	30
<i>Table 11: Constructs Means Based on Ship Class</i>	31
<i>Table 12: Constructs of Praise and Concern by Officer vs Enlisted</i>	37
<i>Table 13: Constructs with Significant Differences in Means Based on Officer vs Enlisted</i>	38
<i>Table 14: Construct Means by Paygrade</i>	38
<i>Table 15: Ship Department Areas of Praise and Concern</i>	45
<i>Table 16: Major Mean Differences between Engineering and Non-Engineering Departments</i>	45
<i>Table 17: Constructs Means and Tests of Significance based on Ship Departments</i>	46
<i>Table 21: Mean Differences Between 1(Basic) and 5 (Sustainment)</i>	65
<i>Table 23: Results of Cluster Analysis</i>	68
<i>Table 24: K-mean Clusters Aggregated into Topics</i>	69
<i>Table 25: Cluster Analysis by Paygrade</i>	69
<i>Table 26: Cluster Analysis by Ship Class</i>	70
<i>Table 27: Ship Type Areas of Praise and Concern by Ship Class</i>	74
<i>Table 28: Tests of Significance and Effect Size for Sleep Disturbance by Ship Class</i>	74
<i>Table 29: Tests of Significance and Effect Size for Sleep Impairment by Ship Class</i>	75
<i>Table 30: Tests of Significance and Effect Size for Sleep Factors by Ship Class</i>	75
<i>Table 31: Tests of Significance and Effect Size for Job Stress by Ship Class</i>	75
<i>Table 32: Tests of Significance and Effect Size for Job Demands by Ship Class</i>	76
<i>Table 33: Tests of Significance and Effect Size for Mental Health by Ship Class</i>	76
<i>Table 34: Tests of Significance and Effect Size for General Health by Ship Class</i>	76
<i>Table 35: Tests of Significance and Effect Size for Physical Health by Ship Class</i>	77
<i>Table 36: Means with Largest Differences based on Officer and Enlisted and Tests of Significance</i>	78
<i>Table 37: Tests of Significance and Effect Size for Sleep Disturbance by Paygrade</i>	79
<i>Table 38: Tests of Significance and Effect Size for Sleep Factors by Paygrade</i>	79
<i>Table 39: Tests of Significance and Effect Size for Sleep Impairment by Paygrade</i>	79
<i>Table 40: Tests of Significance and Effect Size for Job Stress by Paygrade</i>	79
<i>Table 41: Tests of Significance and Effect Size for Job Demands by Paygrade</i>	80



<i>Table 42: Tests of Significance and Effect Size for Mental Health by Paygrade</i>	<i>80</i>
<i>Table 43: Tests of Significance and Effect Size for General Health by Paygrade.....</i>	<i>80</i>
<i>Table 44: Tests of Significance and Effect Size for Physical Health by Paygrade</i>	<i>80</i>



LIST OF FIGURES

Figure 1: Mean Constructs of Interest by Age on Likert Scale, lower number is less favorable	23
Figure 2: Mean Constructs of Interest by Age on Likert Scale, higher number is less favorable	23
Figure 3: Mean Constructs of Interest by Age on Days Scale, lower number is less favorable.....	23
Figure 4: Mean Constructs of Interest by Age on Days Scale, higher number is less favorable	24
Figure 5: Average Hours Spent Per Activity Per Day (SURFOR)	26
Figure 6: Average Hours Spent Per Activity Per Day (CNSP/L)	29
Figure 7: Bar Chart of Sleep Disturbance by Ship Class	31
Figure 8: Bar Chart of Sleep Factors by Ship Class	32
Figure 9: Bar Chart of Sleep Impairment by Ship Class	32
Figure 10: Bar Chart of Job Stress by Ship Class	33
Figure 11: Bar Chart of Job Demands by Ship Class.....	33
Figure 12: Bar Chart of General Health by Ship Class	34
Figure 13: Bar Chart of Mental Health by Ship Class	34
Figure 14: Bar Chart of Physical Health by Ship Class.....	35
Figure 15: Heatmap of Sleep Factors by Ship Class	36
Figure 16: Heatmap of Sleep Constructs by Ship Class	36
Figure 17: Bar Chart of Sleep Disturbance by Paygrade	39
Figure 18: Bar Chart of Sleep Factors by Paygrade.....	39
Figure 19: Bar Chart of Sleep Impairment by Paygrade.....	40
Figure 20: Bar Chart of Job Stress by Paygrade.....	40
Figure 21: Bar Chart of Job Demands by Paygrade	41
Figure 22: Bar Chart of General Health by Paygrade	41
Figure 23: Bar Chart of Mental Health by Paygrade.....	42
Figure 24: Bar Chart of Physical Health by Paygrade	42
Figure 25: DNA Chart of Sleep Factors by Paygrade.....	43
Figure 26: DNA Chart of Sleep Constructs by Paygrade.....	43
Figure 27: Average Hours of Spent per Activity per Day (Paygrade).....	44
Figure 28: Sleep Factors by Ship Department.....	46
Figure 29: Sleep Constructs by Ship Department.....	47
Figure 30: Average Hours Spent per Activity per Day (Department)	47
Figure 31 Comparison of Job Stress Across Different Fit and Fill Criterion Thresholds.	51
Figure 32 Comparison of Job Demands Across Different Fit and Fill Criterion Thresholds.....	52
Figure 33 Comparison of Sleep Obtained Across Different Fit and Fill Criterion Thresholds	53
Figure 34 Comparison of Sleep Deficiency across different Fit and Fill Criterion Thresholds	54
Figure 35 Comparison of Workload Across Different Fit and Fill Criterion Thresholds.....	55
Figure 36 Comparison of Fatigue Impairment Across Different Fit and Fill Criterion Thresholds.....	56
Figure 37 Comparison of Sleep Disturbance Across Different Fit and Fill Criterion Thresholds	57



<i>Figure 38 Comparison of Refreshing Sleep Across Different Fit and Fill Criterion Threshold</i>	58
<i>Figure 39 Comparison of Exercise Time Across Different Fit and Fill Criterion Thresholds</i>	59
<i>Figure 40 Comparison of Leadership Assessment Across Different Fit and Fill Criterion Thresholds</i>	60
<i>Figure 41 Comparison of Affective Commitment Across Different Fit and Fill Criterion Thresholds</i>	61
<i>Figure 42 Comparison of Command Satisfaction Across Different Fit and Fill Criterion Thresholds</i>	62
<i>Figure 43 Comparison of Command Safety Ratings Across Different Fit and Fill Criterion Thresholds</i>	63
<i>Figure 44 Comparison of Social Support Across Different Fit and Fill Criterion Thresholds</i>	64
<i>Figure 45: Bar Chart of Sentiment Score by Paygrade</i>	66
<i>Figure 46: Bar Chart of Sentiment Score by Department</i>	66
<i>Figure 47: Bar Chart Sentiment Score by Ship Class</i>	67
<i>Figure 48: Number of Responses Aligned to Each Cluster by Paygrade</i>	70
<i>Figure 49: Number of Responses Aligned to Each Cluster by Ship Class</i>	71



1 INTRODUCTION

1.1 Executive Summary and Year One Lessons Learned

As year one of the new Afloat Surface Climate Assessment Survey (ASCAS) and the Operational Safety Risk Indicator (OSRI) program efforts come to an end, two categories of important lessons have emerged. First, the survey results themselves are strong evidence for actions and programs to reduce risk and improve human performance. Second, as resources are finite and problems can be insidious and embedded within complex causal systems, proven strategies for gathering and exploiting information will prove invaluable in supporting sound, data-enabled strategic decision-making.

For the ASCAS itself, monitoring, analyzing, and reducing risk is of major importance to fulfill mission requirements and protect human and physical capital. While there are many factors that influence operational safety and performance, human factors are key. Human factors, however, entail complex variables that cannot be directly observed, thus requiring innovative approaches to measure and capture their effects. The ASCAS serves to capture crucial manifestations of otherwise unobservable factors.

The survey has demonstrated that it can empower leadership with valuable data to evaluate key facets of a command's operational safety posture and to better understand the dynamic relationship between overarching safety goals and how each command aligns with these goals. The ASCAS includes several constructs proven to relate to safety outcomes that collectively capture overall safety and help establish a meaningful unit safety profile. These components include measures of an organization's climate perception (e.g., Unit Cohesion, Leadership Assessment, and Job Stress); safety-related measures (e.g., Error Management, Command Safety Practices, and Satisfaction); crew endurance measures (e.g., Sleep Hours, Sleep Disturbance, and General Health); items that capture leader perceptions of subordinates; and open-ended comments that provide opportunities for respondents to share their thoughts and concerns.

For this first annual report, the respondents comprised 33 unique crews (6,459 total responses). Responses were aggregated with other TYCOM data and analyzed to provide insight to the SURFOR leadership. Analysis identified areas of praise and concern, as well as distinctions amongst the various demographics. In the aggregate, some notable findings include:

- FIT/FILL analysis suggests that ships with lower FIT metrics (less than 92%) reported less favorable responses than platforms whose FIT was above 92%, relative to FILL, corresponding to other efforts that show increased operational risk when manning is low. Further analysis is ongoing in conjunction with the SURFMEX effort to define the specific boundaries of risk associated with manpower and manning mismatches.
- Areas of praise centered around Leader Perceptions of Subordinates and Unit Cohesion for both Watch Teams and Work Centers.
- Areas of concern centered around Affective Commitment (mission commitment), Job Stress and Job Demands.
- Sailors aligned to Engineering departments reported less favorable responses for nearly every survey construct in every class of ship.
- Uncomfortable mattresses most negatively impacted Sailors getting good sleep.
- Natural Language Processing (NLP) and machine learning (ML) analyses of open-ended comments found that crew endurance, low manning levels, lack of training, lack of nutritional options, lack of sleep, lack of exercise, high workload, low morale, and mental health concerns were prevalent across all crews, but particularly in the maintenance and basic phases.



SURFOR is well-positioned through the distribution, collection, aggregation, and analysis of the ASCAS to contribute insights into essential human factors and safety risk variables, especially when married to other SURFOR-relevant data systems (e.g., MyNavy Learning, TORIS, CRMD, and the DEOCS).

For the data and data science efforts, the key steps in developing a system for evidence gathering and applied statistical reasoning include:

- **Establishing an Analytics Roadmap.** Developing a data-driven organization is a long-term, incremental, iterative process that requires coordinated efforts by subject-matter experts (SMEs) from multiple domains. Such efforts cover topics ranging from data acquisition, data preparation, multimethodological analysis, output interpretation, and ultimately the delivery of actionable information for leaders at all levels. An analytic roadmap was developed to help to organize the complex interplay of topics and interdisciplinary team members; it has proved invaluable in maintaining direction and monitoring progress.
- **Building an Analytics Platform and Data Pipelines.** Realizing the vision of a data-driven organization requires a robust supporting infrastructure (e.g., an Analytics Platform comprising computing hardware and software) and data pipeline. Historically, analytic products were created using manually intensive processes and siloed or one-off data sources; these processes lead to results which are error prone and difficult to reproduce. The Surface Force (SURFOR) recognized the need for a computational infrastructure and reproducible analytics protocol that could overcome these limitations. OSRI was one of the first use cases within the Jupiter data environment. Jupiter, as the primary analytics platform, provides:
 - a centralized development/production environment that can be shared across organizations.
 - access from any NIPR system, while also offering SIPR capabilities.
 - access to multiple DON datasets and permits the ingestion of new data into the system.
 - an infrastructure centered on SURFOR's analytics that can support best practices in data management and analysis (e.g., reproducibility of results and minimizing redundancy).
 - automation capabilities in the production and delivery of results in standardized dashboards to stakeholders, thus accelerating the delivery of information that could prove invaluable in supporting critical, time-sensitive decisions.

SURFOR analysts are providing results and thought leadership across the Navy in many regards as a result. However, continued expansion and success in new applications is unsustainable unless we invest the time and resources necessary to expand staff and unit capabilities as users and providers of data through education, training, and full participation.

- **Defining data collection and integrity expectations.** Though specific data sources have been identified as promising targets, sharing and extracting robust, trustworthy data has proven challenging. SURFOR addressed some of these issues by developing processes for promoting data integrity (i.e., shadow data) as needed. Though these efforts met the immediate need, they exposed more pressing problems threatening data integrity. Overcoming these challenges requires enforcing best practices across all stages of the data lifecycle and requires coordinated efforts from all stakeholders across all levels of leadership.



- **Establish realistic goals for real results.** Obtaining buy-in from leaders at all levels as to the potential of a robust analytics platform and the products it delivers has proved crucial for the early success of this effort. However, the initial enthusiasm needs to be moderated by an understanding that are significant political/bureaucratic and information technology related challenges that have to be overcome to establish a sustainable effort. In other words, corralling and analyzing multi-source enterprise-wide data is an iterative long-term process (i.e., multiple years). Focusing on immediate, visible outcomes can undermine the long-game. For instance, given the growing staff requests for information, the OSRI team has had to create manual processes to satisfy data calls, creating a false impression as to the ease with which data can be accessed and analyzed.



2 BACKGROUND

2.1 Introduction

Following two major mishaps in 2017, the US Navy recognized the need to change not only its business practices but also to conduct applied human factors research to better monitor and understand the psycho-social components underlying mishaps and near-misses. Within the realm of Surface Forces, human factors research primarily focuses on the capabilities and limitations of human physical and cognitive abilities as well as sociotechnical systems (e.g., organizational structures) composed of personnel subsystems that interact via technology. Given the multiple vulnerabilities in how human beings process information, as well as in how human beings operate complicated technologies, human factors are an essential component in understanding and managing risk.

An emerging consensus in the scientific field highlights the importance of developing effective risk management plans for latent error conditions (i.e., organizational factors that are not actively monitored and therefore go undetected), rather than focusing on error prevention per se. Among many other efforts to address human factors across SURFOR, the ASCAS was established as a requirement within CNSP/CNSLINST 3502.7A. The ASCAS draws from published peer-reviewed scientific studies and is intended to capture the factors that can affect operational safety outcomes, such as latent mishap factors, including crew endurance issues.

2.2 What is Safety Climate?

Safety culture refers to the long-standing attitudes, beliefs, and values held within an organization at the enterprise-level, while safety climate reflects the shared individual perceptions of an organization, based on experience (this can extend from the Navy at large to the Watch Teams) and is associated with perceptions of day-to-day workplace realities. The safety climate focus on day-to-day realities is crucial for capturing the relationship, and possible dissociation, between what is espoused (e.g., safety culture policies, procedures, and practices) and what is actually practiced (e.g., local leadership values and behaviors, communication practices, and training). These perceptions can directly translate into the implementation and upholding of local workplace safety procedures, or not. In other words, safety climate shapes the behaviors and experiences of organizational members and thus affects safety performance and outcomes. Safety climate indicators can predict potential incidents, and they are also associated with Sailors' motivation to engage in safety-related behaviors (e.g., intent to comply with rules or voice safety-related concerns). As such, enhancing understanding a command's safety climate is a promising route for preventing accidents and other negative safety-related outcomes.

An organization's safety climate encompasses several factors, such as safety-oriented managerial values, communication, training, and system operation. Organizational safety climate assessments provide insight into general organizational behaviors, but more importantly, they help to uncover the workplace mindsets and organizational dynamics that influence safety-related behaviors. The annual ASCAS provides the data elements necessary for developing a comprehensive safety climate profile, which ultimately will help to identify priority targets (i.e., threats to operational safety) for intervention and mitigation.



2.3 What is Crew Endurance?

In its simplest form, crew endurance reflects the ability of personnel to maintain optimal performance (within an accepted target range), in the face of physiological, psychological, and environmental stressors. Not unlike astronauts, Sailors must operate within different kinds of environments: those that are *Isolated* (i.e., physically separated from civilization and subject to limited communications and social support), *Confined* (i.e., constrained physically, such as habitat-limited privacy, workspace, social interactions, and recreational activities), and *Extreme* (i.e., geophysical environment, shipboard industrial-like setting, and operational tempo). These Isolated, Confined, and Extreme (ICE) environments can compromise mission preparation and execution, increasing the chances that problematic cognitive or behavioral issues will adversely affect crew endurance, performance, and health, both physical and mental.

Effectively managing ICE environments requires vigilance: all personnel are responsible for protecting against insidious factors whose influence can accumulate and imperceptibly undermine crew endurance (e.g., sleep interruptions, lack of social engagement, lack of physical activity, task-induced fatigue). When mission requirements make it impossible to avoid problematic physiological and psychological environmental stressors, leaders must be able to recognize this, then implement effective and validated countermeasures to help protect mission readiness.

Regardless of the mission or operational realities, any number of causes, known and unknown, can reduce a Sailor's performance. Detecting the presence of those primary factors that can be measured in principle is the first step in developing an effective mitigation system. The factors undermining crew endurance are multifaceted, and their interplay is complex and dynamic. To counteract their influence, leadership must strive to ensure that each Sailor's needs are addressed in a comprehensive, rather than in an isolated, piecemeal way (e.g., believing that drinking Red Bull to counteract fatigue is an effective long-term solution). Such efforts are critical for proactive maintenance of a ship's most important resource, the human system.



3 SURVEY INSTRUMENT

3.1 Overview

The ASCAS draws from multiple scientifically validated resources and was approved for use by the Institutional Review Board at Naval Medical Center Portsmouth, ensuring that the survey adheres to all DoD human subject protections. All participation in the effort was voluntary and responses were solicited anonymously. As presented in the tables, each hypothesized, unobservable, latent construct has been captured with some number of individual questions -- the manifest observable variables. These individual items reflect the survey construct; for example, the *Unit Cohesion* construct is hypothesized to comprise three separate items, each reflecting a different facet of unit cohesion (i.e., Can the members of the unit depend on one another? Are they able to cooperate? Do they stand up for each other?). This multi-item approach to measurement enhances the reliability of the measures (i.e., the consistency of responses across different forms of the same construct) and their content validity (i.e., the extent to which the items faithfully capture the content of the latent construct). When reported, the individual items that constitute the representation of a latent construct are aggregated to produce a single score (i.e., the arithmetic mean).

3.2 Survey Measures

3.2.1 DEMOGRAPHICS

Demographic information gathered by the ASCAS included paygrade (Junior Sailors, Petty Officers, Chiefs and Officers/CWOs), gender (Male and Female), age range (17-21, 22-25, 26-30, 31-35, 36-40, 41-45, 46-50, and 51-55), race (Non-white and White; note that normal categories were not used as to help better ensure anonymity; e.g., to protect the only female CWO in Engineering), years of service (numerical response), months onboard (numerical response), departmental assignment (Admin, AIMD, Air, Combat Systems, Deck, Engineering, Medical/Dental, Operations, Safety, Supply, Training / Plans & Tactics, and Weapons), and Ship Classes (LHD, LHA, LPD, CG, DDG, LCS Freedom, LCS Independence, and Patrol Craft).

3.2.2 ORGANIZATIONS CLIMATE PERCEPTION MEASURES

The primary measures have been validated in peer-reviewed scientific publications and relate directly to operational safety and human performance outcomes. These measures, and brief explanations for their use, include:

Unit Cohesion. Previous military studies have found that unit cohesion can impact the resilience and well-being of personnel. While the word “unit” implies a larger command structure in the military context, a unit in the truest sense reflects a composition of two or more people. Respondents are asked to rate their perceptions of their Watch Team, if applicable, and Work Center unit cohesion using an established 3-item Likert-like scale measure. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Leadership Assessment. Numerous studies have demonstrated that perceptions of unit leadership impact morale, performance, and health outcomes. Consistent with those studies, respondents are asked to rate their perceptions of unit leadership with a 4-item Likert-like scale assessing perceptions of unit leadership behavior. The response format of each item comprised frequency estimates and ranged from 1 (*never*) to 5 (*very often*).



Job Stress. Chronic stress carries with it far-reaching interrelated negative externalities; for example, chronic stress reduces slow-wave sleep (i.e., the good sleep – the only time when the blood-brain barrier increases in permeability and the brain glymphatic system removes metabolic waste products accumulated during normal activity) and this in turn affects cognitive processes which undermine one’s ability to make higher-level decisions (e.g., the capacity to make healthy food choices) and negatively affects one’s mood. Respondents are asked to rate their perceptions of job stress on an established 3-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Affective Commitment. An employee’s emotional attachment to, identification with, and involvement in an organization reflects their affective commitment to the organization, which in turn has implications for an organization’s success. Affective commitment is assessed with an established 3-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Social Support. In general terms, social support reflects the degree to which one has people available whom they can rely upon to provide help/counsel. The lack of social support is associated with increased mental/physical health outcomes as well as decreased worker performance. Respondents are asked to rate their perceptions of social support with an established 4-item Likert-like scale. The response format of each item comprised estimates of how often a particular social support component applied to the current command and ranged from 1 (*none of the time*) to 5 (*all of the time*).

Team Process. An overall team process is defined as “interdependent acts that convert inputs to outcomes through cognitive, verbal, and behavioral activities directed toward organizing taskwork to achieve collective goals. Centrally, team process involves members interacting with other members and their task. They are the means by which members work interdependently to utilize various resources such as expertise, equipment, money, to yield meaningful outcomes (e.g., product development, rate of work, team commitment, satisfaction).” The team process measure is comprised of *transition*, *action*, and *interpersonal* subprocesses. Transition processes occur prior to or between performance episodes and have a dual focus whereby members reflect on and interpret previous accomplishments as well as prepare for future actions. Action processes describe the behaviors that members engage in while working toward goal accomplishment. Interpersonal processes focus on the personal relationships between members. Respondents are asked to rate their perceptions of the command’s team processes with an established 11-item Likert-like scale measure. The response format of each item comprised estimates of the extent to which the command committed to each team process component and ranged from 1 (*not at all*) to 5 (*to a very great extent*).

Crew Resilience. Crew resilience reflects the ship’s ability to adapt to both incremental changes and sudden disruptions in order to effectively satisfy operational demands; the ability to do so is paramount for high-risk organizations. Respondents are asked to rate their perceptions of the crew’s ability to bounce back and thrive in the face of adversity using an established 3-item Likert-like scale measure. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).



3.2.3 SAFETY-RELATED MEASURES

Error Management. The error management construct rests, in part, on the assumption that human errors can never be entirely prevented and therefore focuses on how to mitigate the consequences of errors that have occurred. Whereas error prevention seeks to avoid errors altogether, error management focuses on reducing negative error consequences and amplifying positive consequences (e.g., using the incident as a teaching moment). In sum, error management ensures that errors are quickly reported, negative error consequences are effectively handled/minimized, and that organizational learning transpires. Respondents are asked to rate their perception of the command's error management climate with an established 2-item Likert scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

No-blame Error Reporting. Known issues and mistakes often do not get reported for fear of reprisals. As such, minor issues requiring relatively simple fixes can manifest into major mishaps. A no-blame error reporting atmosphere mediates the relationship between safety climate and negative (e.g., unplanned maintenance) and positive (e.g., improved error detection through proactive crew involvement) outcomes. Respondents are asked to rate their perception of the command's no-blame reporting climate with an established 4-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Psychological Safety. The concept of psychological safety describes one's perception of the consequences associated with taking interpersonal risks (e.g., voicing a concern, sharing an opinion, submitting a recommendation, or taking preemptive action). A large body of scientific research has demonstrated that increased psychological safety enhances a worker's willingness to communicate concerns, or promote organizational knowledge acquisition, employee well-being, and overall performance. Psychological safety is assessed with an established 7-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Caring. Research concerning organizational climate has repeatedly demonstrated that distinctive group-level climates emerge within organizations and locally influence performance outcomes (e.g., work quality and safety behaviors); for example, an officer who directs Sailors to disregard certain safety procedures whenever their department falls behind schedule creates a distinction between Navy procedures and subordinate group-level practices, thus creating the potential for distinctive climates within an organization. Additionally, workers tend to perform better and experience more positive health outcomes when they feel that their organization cares about them. Within the military context, various organization levels can influence Sailor perceptions of care. Respondents thus are asked to rate the degree to which they perceived that the Navy, their command, their department, and their direct supervisor care about their health, safety, and well-being using a 4-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Command Safety Practices. It stands to reason that a command's safety practices factors into operational safety outcomes. Indeed, the scientific literature provides empirical support for this proposition; for example, research in the commercial maritime sector found that worker's perceptions of management safety practices is associated with lower accident and crew fatality rates. Respondents are asked to rate their perceptions of command safety practices with an established 4-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).



Noncompliance. A number of factors contribute to non-compliant behavior in the workplace; and if we were able to understand each Sailor's personal issues and how they see the world then each act of non-compliance might make sense on its own. Unfortunately, a large and high-risk organization like the Navy does not afford such an assessment capability. Nonetheless, it is important to gauge attitudes towards non-compliance in assessing an organization's safety climate. Command-level noncompliance is assessed with an established 3-item Likert-like scale. Individual-level noncompliance is assessed with an established 4-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Satisfaction. A worker's satisfaction with workplace safety practices/policies is positively associated with an organization's overall safety climate, which in turn increases commitment to organizational safety policies and reduces mishaps/accidents. Respondents are asked to rate their level of satisfaction with the command's safety practices with an established 2-item Likert-like scale. Additionally, respondents are asked to rate their level of agreement in regard to both being motivated to serve and being assigned to the command using a 2-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Personal Responsibility for Workplace Safety. Workers often feel a lack of ownership in regard to their work environment which can result in underappreciating the important role that they personally play in the organization's successes and failures. Respondents are asked to rate their perceptions of personal responsibility in regard to command safety with a 4-item Likert-like scale measure. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Safety Communication. One of the most effective ways to promote a positive safety climate and mitigate negative outcomes is to ensure that safety-related comms-traffic across the entire organization is easily facilitated. Additionally, it is important to recognize that each Sailor has a different style of communication (e.g., passive, aggressive, passive-aggressive, and assertive) and you should strive to facilitate input accordingly. Respondents are asked to rate their perceptions of safety communications within the command using an established 3-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Safety Reporting Behaviors. The US Occupational Safety and Health Administration has estimated half of all workplace safety issues/accidents go unreported. A number of reasons underlie such underreporting (e.g., fear of getting reprimanded, a *suck it up* mentality, or just the absence of a system to make either anonymous or non-anonymous reports). Respondents are asked to rate their perceptions of reporting safety issues within the command using an established 3-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

Job Demands. Workplace job demands present individuals with a number of stressors, of varying potential severity. Excessive job demands can exhaust a Sailor's mental and physical resources and ultimately undermine their personal performance as well as that of the entire command. Respondents are asked to rate their perceptions of job demands using an established 5-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*never*) to 5 (*very often*).

Job Resources. Worker perceptions of available job resources help them properly conduct their duties as well as offset workplace pressures, both mentally and physically. Positive perceptions of job resources promote active Sailor workplace engagement and general job satisfaction. Respondents are asked to rate their perceptions of safety-related training and equipment/facilities resources using an established 6-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).



Safety Issues Reporting. Both in the civilian sector and military, potential safety issues and accidents are often not reported. To assess if potential safety issues are under-reported, respondents are asked to indicate the number of times they reported a safety issue to their supervisor, a safety issue went unreported to their supervisor, and they experienced or witnessed a near-miss. The response format of each item comprised levels of agreement and ranged from 1 (*strongly disagree*) to 5 (*strongly agree*).

3.2.4 CREW ENDURANCE MEASURES

Sleep Hours. Problematic sleep impacts human performance by way of functional impairment and can lead to the onset of numerous health issues. Respondents are asked to indicate the number of sleep hours available onboard, number of sleep hours required to feel well-rested, and the number of sleep hours generally obtained onboard and at home.

Naps. There is strong evidence in the scientific literature that naps, less than 30 minutes in duration, promote human performance outcomes (e.g., alertness and cognitive processes). Respondents are asked to indicate how many days per week they take naps, the average duration of their naps, and well-rested they feel following their naps both while onboard the ship and at home.

Sleep Disturbance. Although the amount of sleep obtained each day is important, so is the quality of the sleep. Generally speaking, good quality sleep means that one typically falls asleep within 30 minutes, sleeps soundly with minimal awakenings, and drifts back to sleep within a few minutes if awoken; while bad quality sleep entails one having trouble falling asleep, trouble staying asleep, restlessness, and early terminal awakening (awakening before desired wake time). Sleep quality is assessed with an established 5-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*never*) to 5 (*very often*).

Sleep Functional Impairment. Respondents were asked to rate how much sleepiness or fatigue affects their duties or ability to function as well as falling asleep while on duty. The response format of each item comprised levels of agreement and ranged from 1 (*never*) to 5 (*very often*).

Sleep Factors. General shipboard factors known to negatively impact sleep (i.e., Ambient Noise, 1MC Announcements, Crew Mate Noise, Uncomfortable Mattress, Temperature, Invasive Lighting, Workload, Meetings, Inspections, Drills, Rack Design) are assessed with a 5-item Likert-like scale. The response format of each item comprised levels of agreement and ranged from 1 (*not at all*) to 5 (*extremely*).

Daily Activities. To account for how personnel allocate their wake-time while onboard, respondents were asked to report how they spend their time per day across eight areas (Work Center Duties, Watch Team Duties, Completing Personnel Qualification Standards, Attending Meetings, Exercising Socializing, Personal Time and Other).

General Health. Respondents are asked to rate their overall health on a 1-item Likert-like scale. Additionally, respondents are asked to indicate how many days their mental and physical health was not good in the past 30 days.



Physical Fitness Activity. Physical activity provides a number of benefits beyond simply enhancing one's ability to pass a PFT. Physical activity affects brain plasticity and improves cognitive performance (e.g., visual pattern recognition), locomotion (e.g., balance) and mood enhancer. Conversely, a notable decrease in physical activity is a possible indicator of depression. To assess shipboard physical activity, respondents are asked to indicate how many days per week they got vigorous/high-intensity exercise (i.e., experience increased heart rate, breathing rate, sweating, and muscle fatigue) while underway and in-port. Respondents are asked to rate the adequacy of the shipboard fitness facilities/equipment and time allotted to exercise with a 2-item Likert-like scale measure. Additionally, respondents are asked to rate to indicate whether they are required to partake in routine command PT as to gauge the command's commitment to physical fitness.

General Diet. Proper nutrition is a key component to maximize human performance, weight management (via energy balance) and improve overall health by reducing disease susceptibility. In relation to crew endurance, one's diet influences sleep quality, general energy levels, and human performance by way of macro and micro-nutrient composition. Respondents are asked to indicate how healthful (i.e., nutritious and comprised of proper serving sizes) their overall diet while both at home and onboard the ship.

Sun Exposure. While excessive sunlight exposure can result in skin cancer, natural sunlight has important biological function for humans. Namely, sun exposure helps inform the biological circadian (near 24-hour) clock by which the timing of every physiological function in the body is regulated. Secondly, natural sunlight exposure helps the body produce Vitamin D, which serves a key role in sleep processes, bone health, and helps combat negative health outcomes (e.g., inflammation, high blood pressure, muscle fatigue). Additionally, natural sunlight exposure elicits biological responses (e.g., serotonin production) that counter the onset mental health issues (e.g., depression and seasonal affective disorder) while enhancing cognitive performance. Respondents are asked to rate to indicate how many minutes of sun exposure they generally receive while underway with a 1-item measure.

Caffeine Usage. While caffeine can enhance some aspects of human performance, it is a psychoactive drug (i.e., impacts mental processes) which can also undermine performance. Excessive caffeine consumption has become an area of concern in both the general and military populations with the rapid expansion of energy drinks, supplements, and stimulants in the marketplace, to include ship stores. Generally speaking, consumption of approximately 400 milligrams of caffeine (about four 8-ounce cups of standard brewed coffee) per day appears to be safe for most healthy adults, while 250 milligrams of caffeine per day is considered moderate usage and 500 milligrams is considered excessive usage. Concerning ingredients aside, the mass consumption of energy drinks/shots is presenting new concerns as the amount of caffeine per serving size ranges from 35 to >300 milligrams. Energy drinks/shots consumption cannot only disrupt sleep, but also cause cardiovascular system, neurological, gastrointestinal, renal, and endocrine system problems. Caffeine can adversely impact general crew performance by disrupting their circadian cycle as it blocks sleep-inducing chemicals in the brain and increases adrenaline production, which in turn destabilizes cognitive, emotional, and behavioral processes. Respondents are asked how many times per day they consume various sources of caffeine.

Nicotine Usage. Nicotine is a stimulant that can disrupt one's circadian cycle. Additionally, most nicotine delivery systems, whether it be via an e-cigarette or traditional cigarette, expose the user to harmful chemicals. Smoking is also associated with an increased prevalence of sleep-related respiratory disorders, which further undermines sleep quality and causes daytime sleepiness. Respondents are asked how many times per day they consume various sources of nicotine.



3.2.5 LEADER PERCEPTIONS OF SUBORDINATES

Procedural Compliance. Although the onus of organizational outcomes generally falls upon Commanding Officers, it is recognized in the academic literature that subordinates also carry a share of the responsibility. Unfortunately, DoD assessment tools such as the DEOCS, often solely focus on Commanding Officers. To provide a new perspective, an established measure is used to assess leader perceptions of subordinates (i.e., those who supervise at least 2 Sailors) with respect to workplace compliance, risk management, and general safety behaviors. The response format of each item comprised levels of agreement and ranged from 1 (*never*) to 5 (*very often*).

3.2.6 OPEN-ENDED COMMENTS

Respondents were given the opportunity to share any thoughts that they might have concerning safety-related issues (e.g., methods to improve workplace safety or current safety-related issues), best/worst safety practices, and/or crew endurance issues (e.g., sleep, exercise, and nutrition).



4 RESULTS

4.1 Introduction

This report used exploratory data analysis techniques to evaluate survey results. Descriptive statistics, including means and frequency tables, were applied. Analysis of variance (ANOVA) techniques were used to compare mean differences in groups across continuous variables. Categorical variables include paygrade, ship department, and ship type. Group means and differences in means are provided, along with standardized mean differences as a measure of effect size (a difference in means rescaled by pooled variance, Cohen's *d*). While there is no definitive cutoff for interpreting Cohen's *d*, benchmarks often used include 0.2, 0.5, and 0.8 as small, medium, and large effects, respectively. Additionally, Natural Language Processing (NLP) and Machine Learning (ML) techniques were applied to quantify, categorize, and derive insights into the open-ended questions. Large sample sizes, such as these, can produce statistically significant outcomes, even for trivial differences in size; as such, any use of a *p*-value should thus be interpreted with caution, and, at the very least, combined with measures of the effect size.

4.2 COMNAVSURFOR (Overall)

A total of 33 crews completed the ASCAS survey in its first year. Although the ASCAS is a mandatory crew requirement, completion of the survey by the crewmembers is completely voluntary. While there were 6,459 respondents, just 3,313 Sailors opted-in to complete the survey. The analysis below is based on these 3,313 responses, which span all 33 ships.

4.2.1 DEMOGRAPHICS

Demographics of the sample are provided in Table 1, which includes counts of participants based on gender, race, age, paygrade, and the department to which they are assigned. In the ship departments group, the number of participants is provided in Table 1, but due to distinct differences from the other departments, this variable has been dichotomized into Engineering and Non-Engineering groups.



Table 1: Demographics of the ASCAS Respondents (SURFOR)

	CNSF	%	CNSP	%	CNSL	%
Gender						
Female	646	20.26%	438	21.66%	208	17.82%
Male	2543	79.74%	1584	78.34%	959	82.18%
Age						
17-21	474	14.83%	294	14.51%	180	15.38%
22-25	894	27.97%	568	28.04%	326	27.86%
26-30	658	20.59%	409	20.19%	249	21.28%
31-35	587	18.37%	391	19.30%	196	16.75%
36-40	346	10.83%	217	10.71%	129	11.03%
41-45	185	5.79%	118	5.82%	67	5.73%
46-50	42	1.31%	23	1.14%	19	1.62%
51-55	10	0.31%	6	0.30%	4	0.34%
Race						
Non-white	1406	44.79%	913	45.88%	493	42.91%
White	1733	55.21%	1077	54.12%	656	57.09%
Paygrade						
Junior Sailors	474	14.82%	289	14.25%	185	15.80%
Petty Officers	1838	57.46%	1139	56.16%	699	59.69%
Chiefs	363	11.35%	239	11.79%	124	10.59%
Officer/CWO	524	16.38%	361	17.80%	163	13.92%
Ship Department						
Admin	209	6.60%	143	7.13%	66	5.69%
AIMD	72	2.27%	1	0.05%	71	6.12%
Air	66	2.08%	12	0.60%	54	4.66%
Combat Systems	658	20.78%	422	21.03%	236	20.34%
Deck	184	5.81%	116	5.78%	68	5.86%
Engineering	732	23.11%	527	26.26%	205	17.67%
Medical/Dental	89	2.81%	55	2.74%	34	2.93%
Operations	478	15.09%	328	16.34%	150	12.93%
Safety	1	0.03%	0	0.00%	1	0.09%
Supply	289	9.13%	164	8.17%	125	10.78%
Training / Plans & Tactics	51	1.61%	38	1.89%	13	1.12%
Weapons	338	10.67%	201	10.01%	137	11.81%
Ship Class						
Amphib (LHD, LHA, LPD)	1168	35.26%	663	31.02%	505	42.94%
CG	452	13.64%	129	6.04%	323	27.47%
DDG	1395	42.11%	1075	50.30%	320	27.21%
LCS (Free, Ind)	255	7.70%	227	10.62%	28	2.38%
PC	43	1.29%	43	2.01%	0	0.00%

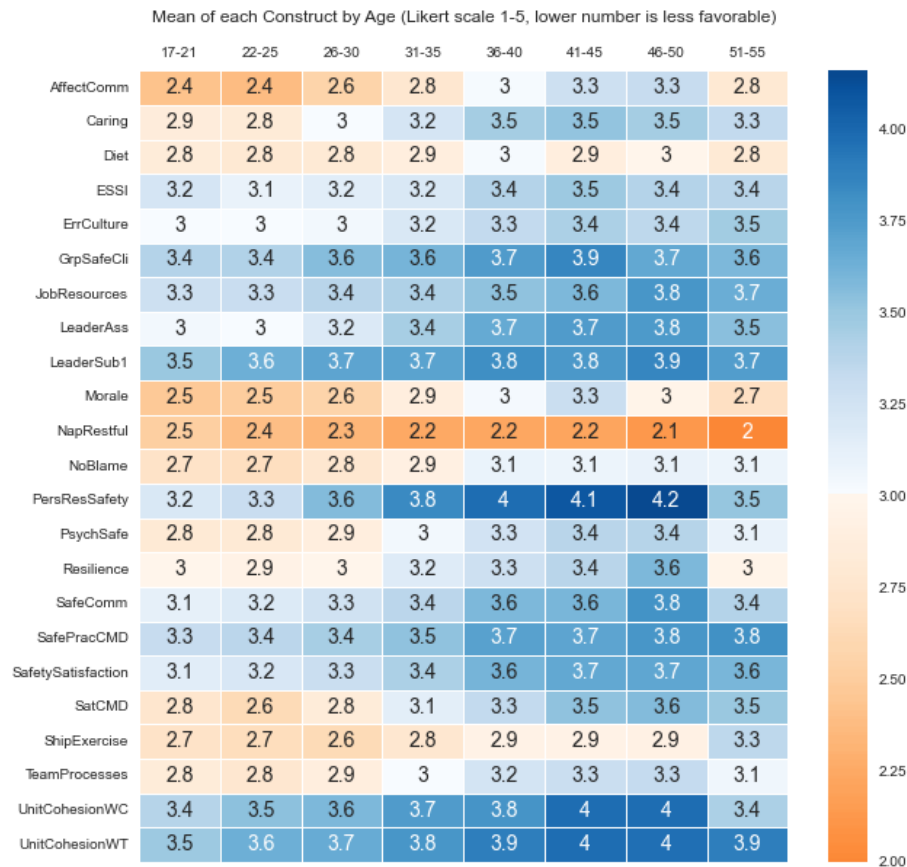


Figure 1: Mean Constructs of Interest by Age on Likert Scale, lower number is less favorable

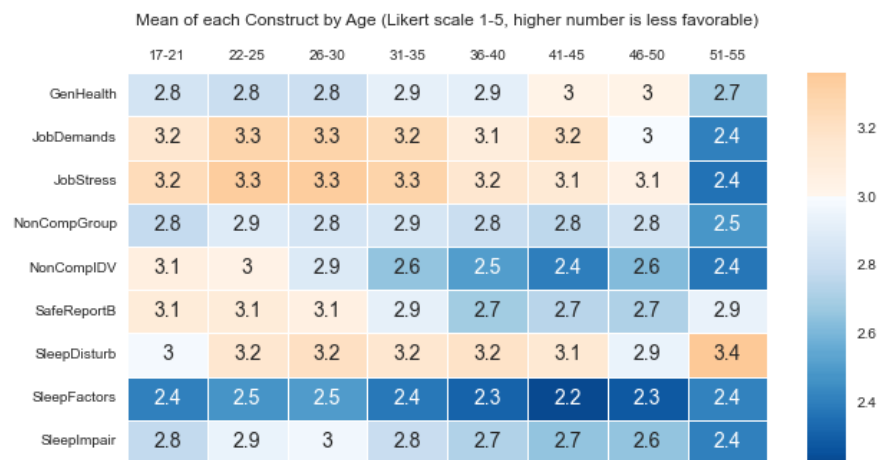


Figure 2: Mean Constructs of Interest by Age on Likert Scale, higher number is less favorable

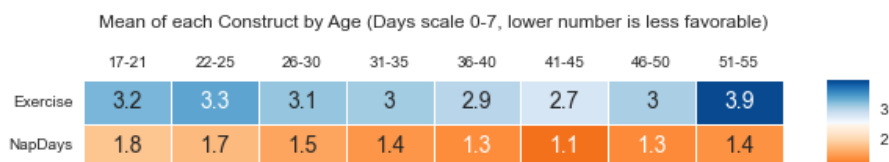


Figure 3: Mean Constructs of Interest by Age on Days Scale, lower number is less favorable

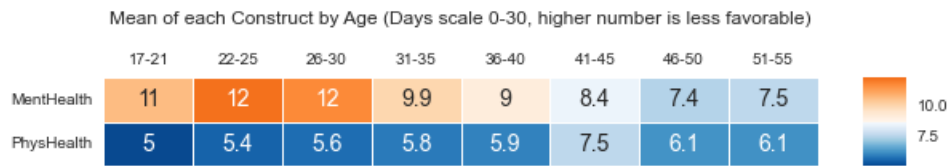


Figure 4: Mean Constructs of Interest by Age on Days Scale, higher number is less favorable

Most respondents were male (79.74%), and the majority were younger than age 35. In terms of racial makeup, 55% of the respondents were categorized as White, with the remainder being categorized as Non-White. The most frequent rank was Petty Officer. These demographics are reflective of the COMNAVSURFOR demographics.

4.2.2 AREAS OF PRAISE AND CONCERN

This section identifies areas of praise and concern for SURFOR as a whole. Respondents are asked to rate their perceptions on a 5-point Likert-scale, with 1 being an unfavorable response and 5 being a favorable response. However, in some instances, questions are reverse encoded such that 1 is a favorable response. Those constructs are denoted with an * in subsequent tables. Areas of praise include constructs that reflect overall favorable responses. Areas of concern indicate negative overall responses. Table 2 identifies high and low scoring constructs for all survey respondents.

Table 2: Areas of Praise and Concern (SURFOR)

Item	Mean (Likert: 1-5)
Leader Perception of Subordinates (Subset 3)	3.89
Leadership Perception of Subordinates (Subset 2)	3.71
Unit Cohesion (Watch Team)	3.69
Unit Cohesion (Work Center)	3.62
Leadership Perception of Subordinates (Subset 1)	3.59
Supervisor Safety	3.58
Personal Responsibility for Workplace Safety	3.57
Command Safety Practices	3.47
Job Resources (Equipment)	3.43
Affective Commitment	2.63
Job Stress *	3.63
No-Blame Error Reporting	2.82
Job Demands *	3.23
Team Process (Interpersonal Processes)	2.85
Psychological Safety	2.91
Satisfaction of Command	2.93
Team Process (Action Processes)	2.93
Team Processes (Overall)	2.93

Notes:

1. For the table above, higher scores represent a more favorable response, with a mean of 3 being neutral. Favorable scores are colored green, and unfavorable scores are shaded red. Scores within 0.05 of neutral (3) are shaded white.

2. * For these questions lower scores represent a more favorable response, with a mean of 3 being neutral.

4.2.3 RESPONSES TO KEY CONSTRUCTS

Portions of the analyses were highlighted due to observed differences between demographics or ship characteristics. Others were determined by prior expectations about a possible relationship to safety, and their effect on preventing or promoting mishaps at sea. These constructs covered aspects of sleep, job demands, job stress, mental, general, and physical health. Tables 3-5 provide the means for the different demographic groups, stratified by the major constructs analyzed.



Table 3: Means of Constructs by Gender

Construct	Gender		Race	
	Female	Male	Non-white	White
Sleep Disturbance *	3.50	3.30	3.35	3.33
Sleep Factors *^	2.51	2.40	2.52	2.35
Sleep Impairment *	2.94	2.82	2.88	2.81
Job Stress *	3.68	3.62	3.54	3.70
Job Demands *	3.24	3.23	3.19	3.26
General Health *^	2.76	2.88	2.82	2.88
Mental Health *+	13.16	10.18	10.71	10.79
Physical Health *+	6.49	5.43	6.12	5.26

Note*: For these questions lower scores represent a more favorable response

Note *+ : For these questions, the values represent the Number of Days. They are also not constructs; therefore, a Cronbach Alpha calculation would not be relevant.

Note *^ : This question is not a construct; therefore, a Cronbach Alpha calculation would not be relevant.

Across all constructs, differences across subgroups of race and gender were generally not pronounced, with the notable exception that women reported more issues with sleep and mental/physical health than men.

Table 4: Means of Constructs by Paygrade

Construct	Paygrade				Department	
	Junior Sailors	Petty Officers	Chiefs	Officer/CWO	Engineering	Other
Sleep Disturbance *	3.23	3.47	3.37	2.98	3.35	3.33
Sleep Factors *^	2.49	2.55	2.20	2.07	2.49	2.41
Sleep Impairment *	2.78	2.96	2.70	2.64	3.02	2.80
Job Stress *	3.42	3.66	3.67	3.70	3.81	3.58
Job Demands *	3.09	3.25	3.25	3.28	3.43	3.17
General Health *^	2.81	2.73	2.93	3.27	2.73	2.89
Mental Health *+	10.50	12.06	9.36	7.57	11.80	10.62
Physical Health *+	5.61	6.00	6.33	4.02	6.89	5.37

Note*: For these questions lower scores represent a more favorable response

Note *+ : For these questions, the values represent the Number of Days. They are also not constructs; therefore, a Cronbach Alpha calculation would not be relevant.

Note *^ : This question is not a construct; therefore, a Cronbach Alpha calculation would not be relevant.

Officers reported better mental and physical health and noted fewer issues with Sleep Disturbances. Junior Sailors reported more mental health issues, Petty Officers reported the highest levels of mental health concerns, with Chiefs following behind. The differences were not as pronounced between engineering and other departments, except that engineering reported slightly worse physical and mental health.



Table 5: Means of Constructs by Age

Construct	Age							
	17-21	22-25	26-30	31-35	36-40	41-45	46-50	51-55
Sleep Disturbance *	3.19	3.36	3.41	3.34	3.37	3.36	3.15	3.65
Sleep Factors *^	2.41	2.48	2.51	2.39	2.32	2.23	2.30	2.36
Sleep Impairment *	2.77	2.94	2.97	2.79	2.73	2.65	2.65	2.40
Job Stress *	3.51	3.69	3.69	3.69	3.52	3.57	3.48	2.67
Job Demands *	3.16	3.29	3.30	3.25	3.09	3.20	2.98	2.40
General Health *^	2.84	2.80	2.81	2.92	2.91	3.03	3.02	2.70
Mental Health *+	10.58	12.21	11.67	9.90	9.01	8.38	7.36	7.50
Physical Health *+	5.02	5.43	5.56	5.76	5.90	7.49	6.14	6.10

Note*: For these questions lower scores represent a more favorable response

Note *+ : For these questions, the values represent the Number of Days. They are also not constructs; therefore, a Cronbach Alpha calculation would not be relevant.

Note *^ : This question is not a construct; therefore, a Cronbach Alpha calculation would not be relevant.

Differences across age groups differed little across most constructs, though a noteworthy exception was the decline in mental health issues and the corresponding increase in physical health issues with age.

4.2.4 ACTIVITY ANALYSIS

A portion of the survey focuses on how Sailors spend their time on a daily basis. Figure 1 shows the average lengths of time individuals report working in the following areas: Work center duties, Watch Team duties, completing personnel qualification standards, and attending meetings.

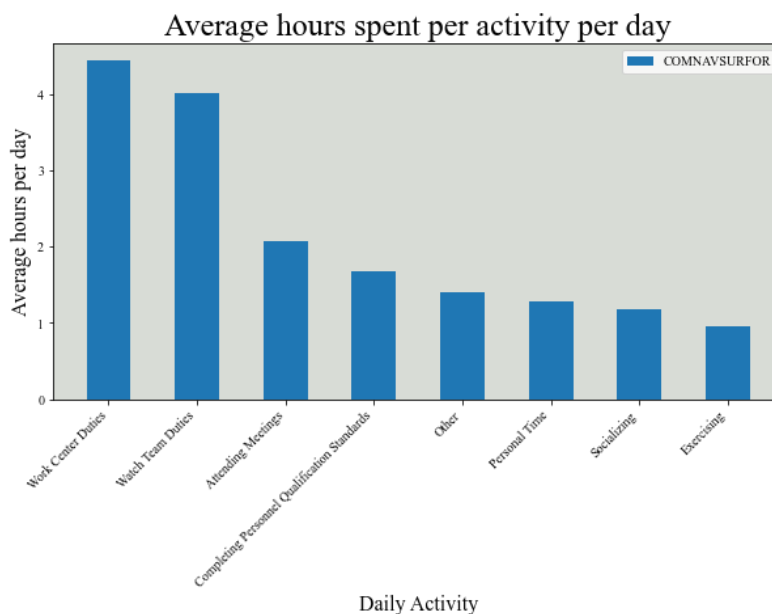


Figure 5: Average Hours Spent Per Activity Per Day (SURFOR)

The data shows respondents spend most of their time (4 hours and 26 minutes) completing Work Center Duties, followed by Watch Team Duties (4 hours) and Attending Meetings (2 hours and 4 minutes). On average, respondents reported spending (12 hours and 12 minutes) hours on work-related duties and (4 hours and 49 minutes) hours on personal activities.



4.3 CNSP/CNSL Analysis

This section explores possible differences between the CNSP and CNSL. In the first year of the new ASCAS survey, 25 crews (2,137 respondents) were aligned to CNSP platforms, and 8 crews (1,176 respondents) were aligned to CNSL platforms.

4.3.1 AREAS OF PRAISE AND CONCERN

Similar to Section 4.2.2, analysis was performed to identify areas of praise and concern. Table 6 highlights these areas for both CNSP and CNSL.

Table 6: CNSP/L Areas of Praise and Concern

Item	CNSP Mean (Likert: 1-5)	CNSL Mean (Likert: 1-5)
Leader Perception of Subordinates (Subset 3)	3.91	3.86
Unit Cohesion (Watch Team)	3.72	3.64
Leadership Perception of Subordinates (Subset 2)	3.72	3.69
Unit Cohesion (Work Center)	3.67	3.54
Supervisor Safety	3.61	3.51
Personal Responsibility for Workplace Safety	3.61	3.50
Leadership Perception of Subordinates (Subset 1)	3.60	3.57
Command Safety Practices	3.52	3.39
Job Resources (Equipment)	3.49	3.34
Job Stress *	3.71	3.50
Affective Commitment	2.69	2.52
Job Demands *	3.29	3.13
No-Blame Error Reporting	2.83	2.81
Team Process (Interpersonal Processes)	2.90	2.76
Psychological Safety	2.95	2.83
Team Process (Action Processes)	2.97	2.85
Team Processes (Overall)	2.97	2.86
Satisfaction of Command	2.99	2.83
Team Process (Transition Processes)	2.97	2.86

Note:

1. For the table above, higher scores represent a more favorable response, with a mean of 3 being neutral. Favorable scores are colored green, and unfavorable scores are shaded red. Scores within 0.05 of neutral (3) are shaded white.

2. *For these questions lower scores represent a more favorable response, with a mean of 3 being neutral.



4.3.2 CONSTRUCTS WITH NOTABLE DIFFERENCES IN MEANS

Mean differences were analyzed between the CNSP and CNSL Table 7 provides the survey constructs with the largest differences in means. Note, in general, most differences are small to negligible.

Table 7: Constructs with Notable Differences in Means

Item	CNSP			CNSL			T-test		
	Mean	SD	Alpha	Mean	SD	Alpha	Mean Difference	P-Value	Effect Size
Job Stress Construct *	3.71	0.92	0.84	3.50	0.96	0.84	0.21	<0.001***	0.223
Occupational Challenges Construct *	2.84	0.99	0.71	2.59	0.98	0.74	0.25	<0.001***	0.254
Affective Commitment Construct	2.69	1.12	0.92	2.52	1.09	0.90	0.17	<0.001***	0.154
Error Management Construct	3.39	0.88	0.69	3.21	0.89	0.72	0.18	<0.001***	0.203
Command Noncompliance Construct *	2.78	0.95	0.91	2.93	0.94	0.91	-0.15	0.425	0.159
Job Demands Construct *	3.29	0.89	0.83	3.13	0.88	0.84	0.16	<0.001***	0.181
Satisfaction with Command Construct	2.99	1.19	0.84	2.83	1.16	0.83	0.16	<0.001***	0.136
Days which mental health was judged 'not good' *	11.38	10.84		9.99	10.54		1.39	<0.001***	0.130
Days which physical health was judged 'not good'*	6.07	8.80		5.06	8.12		1.01	<0.001***	0.119

* For these items lower scores represent a more favorable response, with a mean of 3 being neutral.

The constructs above are those that exhibited the largest differences in their means. Most differences were statistically significant, though these effects were generally small. CNSP expressed slightly more favorable responses towards Affective Commitment, Error Management, Command Noncompliance, and Satisfaction with Command; where CNSL expressed slightly more favorable views toward Job Stress, Occupational Challenges, Job Demands, Days which mental health was judged “not good”, and Days which physical health was judged “not good”.

4.3.3 RESPONSES TO KEY CONSTRUCTS

While there were 26 unique constructs in the ASCAS Survey, 8 were analyzed in greater depth. These constructs were identified by subject matter experts as directly related to operational safety and human performance outcomes. Specifically, the constructs are Sleep Disturbance, Sleep Factors, Sleep Impairment, Job Stress, Job Demands, Mental Health, General Health, Physical Health. Sleep Disturbance asks about the quality of their sleep, Sleep Factors asks what impacts their sleep, Sleep Impairments asks about whether or not they feel fatigued during the day. Job Stress asks about their feelings about what frustrations they have about their job and challenges of their jobs, Job Demands looks at whether or not they have enough time to complete their tasks as well as equipment and manning issues. Mental health asks how many days they report having had mental health issues in past 30 days, General Health asks participants to rate their overall health and Physical Health. Table 8 details construct means and statistical significance with regard to these constructs.



Table 8: Means of Constructs based on Fleet Status and Tests of Significance between the Two Means

Construct	Fleet Status		Mean Difference	P-Value	Effect Size
	CNSP	CNSL			
Sleep Disturbances *	3.32	3.38	-0.06	0.01	-0.05
Sleep Factors *^	2.37	2.51	0.13	<0.01	0.12
Sleep Impairment *	2.90	2.76	-0.14	<0.01	-0.10
Job Stress *	3.71	3.50	0.22	<0.01	0.20
Job Demands *	3.29	3.13	0.16	<0.01	0.13
General Health *^	2.83	2.89	-0.06	0.10	-0.06
Mental Health *+	11.38	9.99	1.39	<0.01	0.13
Physical Health *+	6.07	5.06	1.01	<0.01	0.12

Note*: For these questions lower scores represent a more favorable response

Note *+ : For these questions, the values represent the Number of Days. They are also not constructs; therefore, a Cronbach Alpha calculation would not be relevant.

Note *^ : This question is not a construct; therefore, a Cronbach Alpha calculation would not be relevant.

All constructs, except for General Health, were statistically significantly different between Pacific compared to Atlantic ships. Pacific was higher for Sleep Factors, Job Stress, Job Demands, Mental and Physical Health. Pacific ships were lower for Sleep Impairment and Sleep Disturbance.

4.3.4 ACTIVITY ANALYSIS

Similar to Section 4.2.4, analysis was performed to identify differences in how much respondents spend on various activities. Figure 2 shows the average lengths of time individuals report allotted to various activities for CNSP and CNSL.

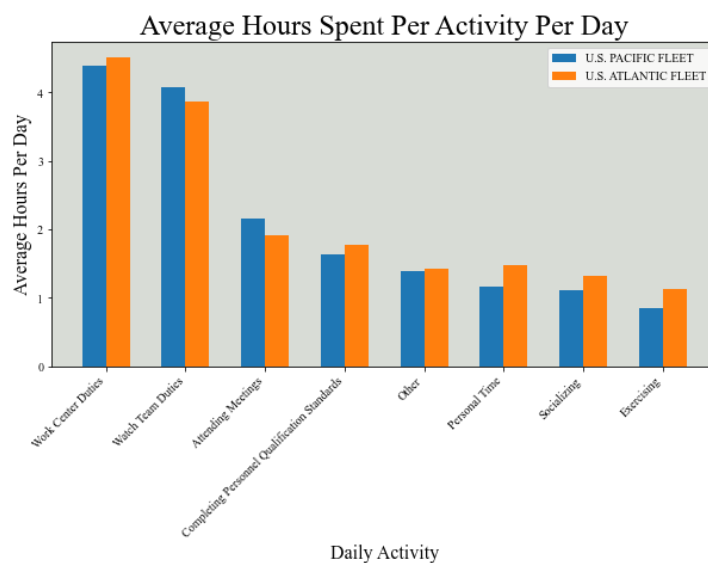


Figure 6: Average Hours Spent Per Activity Per Day (CNSP/L)

CNSP Sailors reported spending an average of 12 hours and 15 minutes on work-related activities and 3 hours and 45 minutes on non-work-related activities. While CNSL Sailors reported spending 12 hours and 5 minutes on work-related activities and 5 hours and 21.5 minutes on non-work-related activities.



4.4 Ship Class Analysis

Results include responses from five main Ship Classes: Amphib Ships (LHD, LHA, LPD), CG, DDG, LCS, PC (MCM). Table 9, below, details the number of crews and respondents aligned to each Ship Class.

Table 9: Number of Ships and Sailors by Ship Class

Ship Class	# of Ships	# Surveyed
Amphib (LHD, LHA, LPD)	7	1168
CG	4	452
DDG	12	1395
LCS	9	255
PC	1	43

4.4.1 AREAS OF PRAISE AND CONCERN

Areas of praise and concern are the constructs with the highest (i.e., praise) and lowest (i.e., concern) means for that category. Table 10 presents areas of praise and concern for each Ship Class.

Table 10: Ship Class Areas of Praise and Concern

Item	Destroyer Mean (Likert: 1-5)	Cruiser Mean (Likert: 1-5)	Amphib Mean (Likert: 1-5)	Littoral Combat Ship Mean (Likert: 1-5)
Leader Perception of Subordinates (Subset 3)	3.97	3.79	3.82	4.09
Leadership Perception of Subordinates (Subset 2)	3.79	3.64	3.62	3.88
Unit Cohesion (Watch Team)	3.75	3.68	3.58	3.97
Unit Cohesion (Work Center)	3.7	3.63	3.45	3.89
Leadership Perception of Subordinates (Subset 1)	3.63	3.58	3.54	3.73
Supervisor Safety	3.6	3.54	3.53	3.75
Personal Responsibility for Workplace Safety	3.57	3.49	3.54	3.88
Command Safety Practices	3.5	3.42	3.48	3.64
Job Resources (Equipment)	3.49	3.22	3.42	3.49
Safety Satisfaction	3.37	3.34	3.31	3.61
Job Stress *	3.72	3.57	3.42	3.79
Affective Commitment	2.72	2.66	2.39	3.01
No-Blame Error Reporting	2.79	2.93	2.76	3.08
Job Demands *	3.28	3.18	3.06	3.31
Team Process (Interpersonal Processes)	2.92	2.9	2.73	3.08
Psychological Safety	2.94	2.95	2.79	3.21
Team Process (Action Processes)	2.98	2.91	2.86	3.17
Safety Reporting Behavior *	3	3.03	2.98	2.78
Team Processes (Overall)	3	2.95	2.84	3.15
Satisfaction of Command	3.02	2.96	2.8	3.11
Team Process (Transition Processes)	3.1	3.07	2.94	3.17
Sleep Disturbances *	3.31	3.38	3.36	3.47

Note:

1. For the table above, higher scores represent a more favorable response, with a mean of 3 being neutral. Favorable scores are colored green, and unfavorable scores are shaded red. Scores within 0.05 of neutral (3) are shaded white.

2. *For these questions lower scores represent a more favorable response, with a mean of 3 being neutral.



4.4.2 CONSTRUCT DEEP DIVE

This section examines the targeted constructs across Ship Class. In Table 11, the overall means for each ship type are provided along with an analysis of the differences for each construct to determine which types are statistically significant for that construct is given in each analysis below.

Table 11: Constructs Means Based on Ship Class

Construct	Ship Class				
	Amphibs (LHD, LHA, LPD)	CG	DDG	LCS	PC
Sleep Disturbance *	3.33	3.38	3.31	3.47	3.20
Sleep Factors *^	2.49	2.51	2.39	2.24	2.19
Sleep Impairment *	2.79	2.84	2.89	2.94	2.80
Job Stress *	3.52	3.57	3.72	3.79	3.73
Job Demands *	3.17	3.18	3.28	3.31	3.14
General Health *^	2.90	2.85	2.82	2.82	2.98
Mental Health *+	10.91	10.42	11.13	10.41	9.58
Physical Health *+	5.82	5.57	5.72	5.24	7.00

Note*: For these questions lower scores represent a more favorable response

Note *+ : For these questions, the values represent the Number of Days. They are also not constructs; therefore, a Cronbach Alpha calculation would not be relevant.

Note *^ : This question is not a construct; therefore, a Cronbach Alpha calculation would not be relevant.

The means above show a relatively stable trend across Ship Class with physical health issues higher on the PC. Interestingly, while the PC reported more days of poor physical health, fewer days of poor mental health were reported.

4.4.2.1 Sleep Disturbance

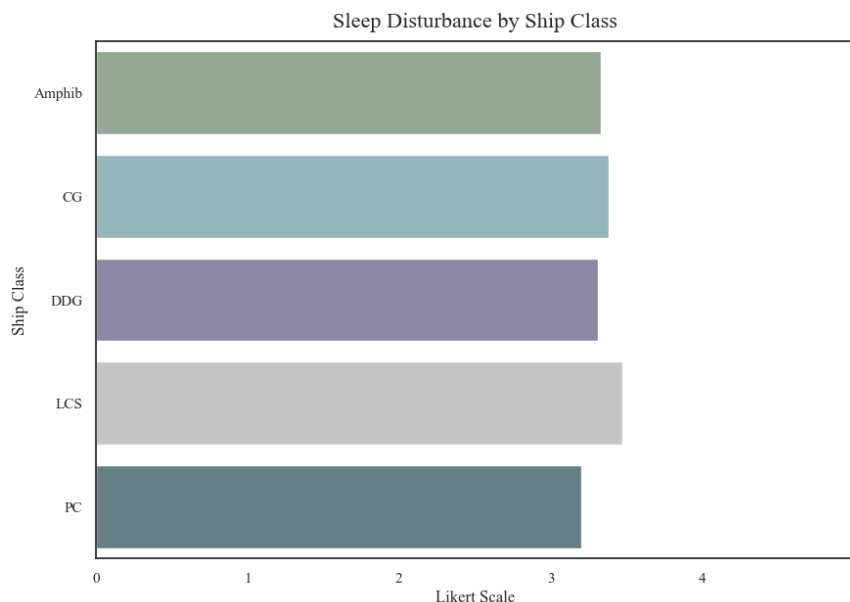


Figure 7: Bar Chart of Sleep Disturbance by Ship Class

LCS and CG ship types reported slightly worse responses for quality of sleep, compared to PC and DDG ship types which reported more favorable responses for quality of sleep.



4.4.2.2 Sleep Factors

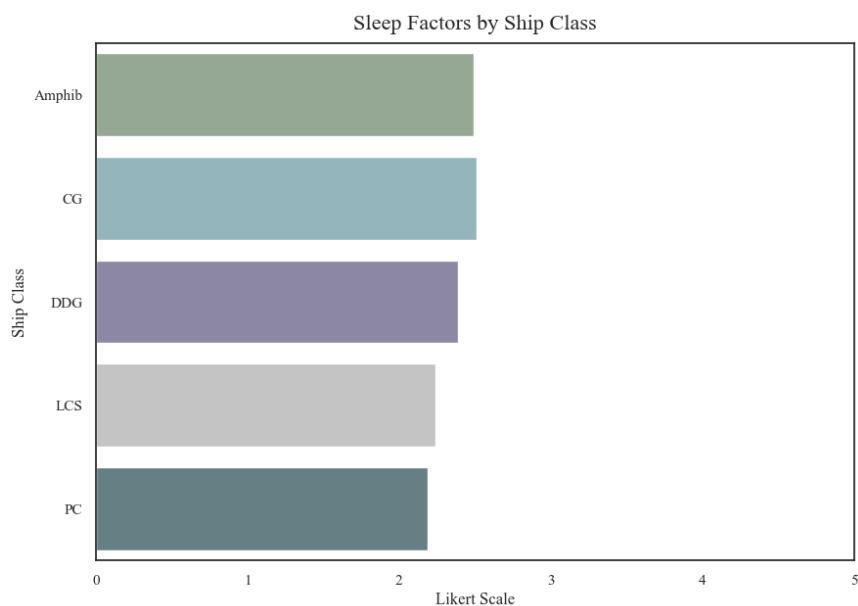


Figure 8: Bar Chart of Sleep Factors by Ship Class

A similar trend was observed across ship classes for Sleep Factors, except for CG and Amphibs which reported more issues impacting sleep compared to others.

4.4.2.3 Sleep Impairment

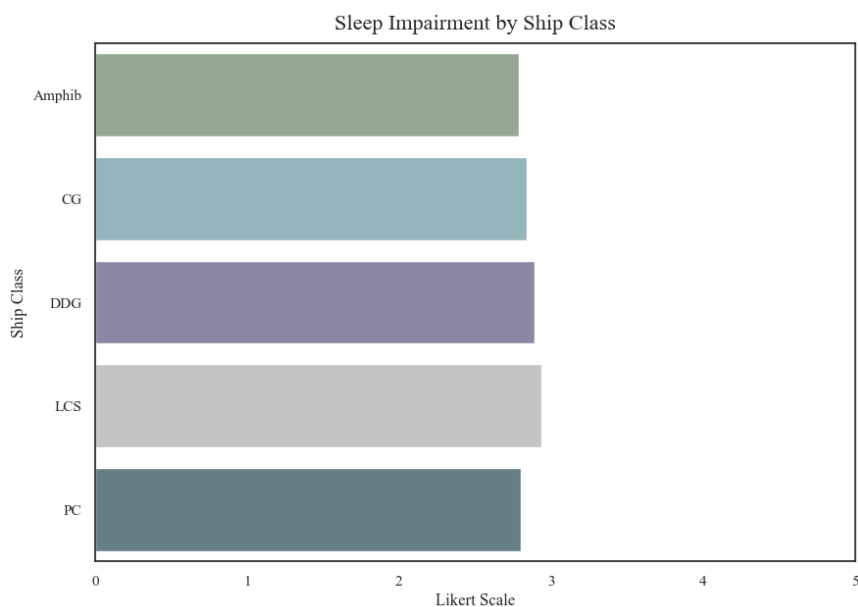


Figure 9: Bar Chart of Sleep Impairment by Ship Class

Sleep Impairments were fairly consistent across all ship classes, LCS and DDG reported slightly more sleep impairments than other ships.



4.4.2.4 Job Stress

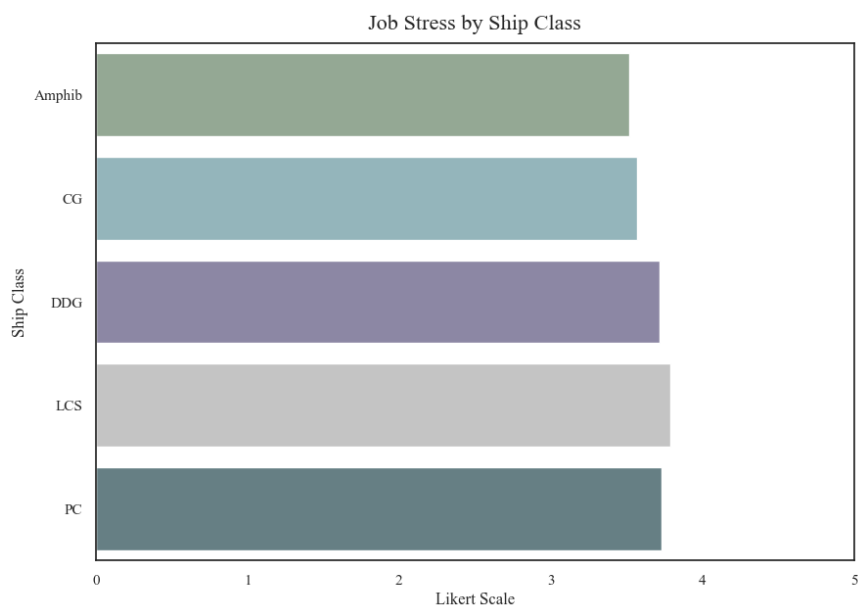


Figure 10: Bar Chart of Job Stress by Ship Class

Differences in Job Stress were not pronounced across ship classes, but LCS and PC reported slightly less favorable responses with regard to job stress.

4.4.2.5 Job Demands

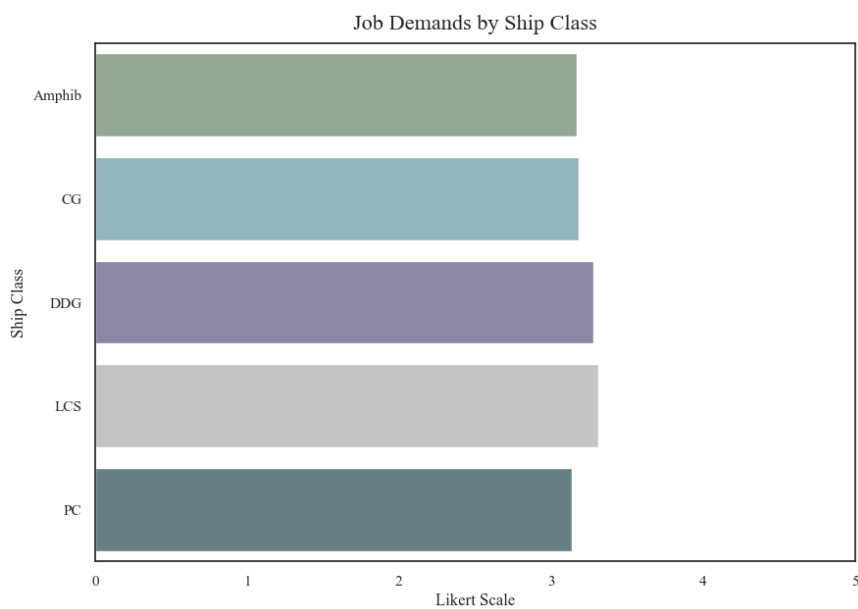


Figure 11: Bar Chart of Job Demands by Ship Class

Although Job Demands had similar responses across ship classes, LCS and DDG reported more job demand stressors compared to others.



4.4.2.6 General Health

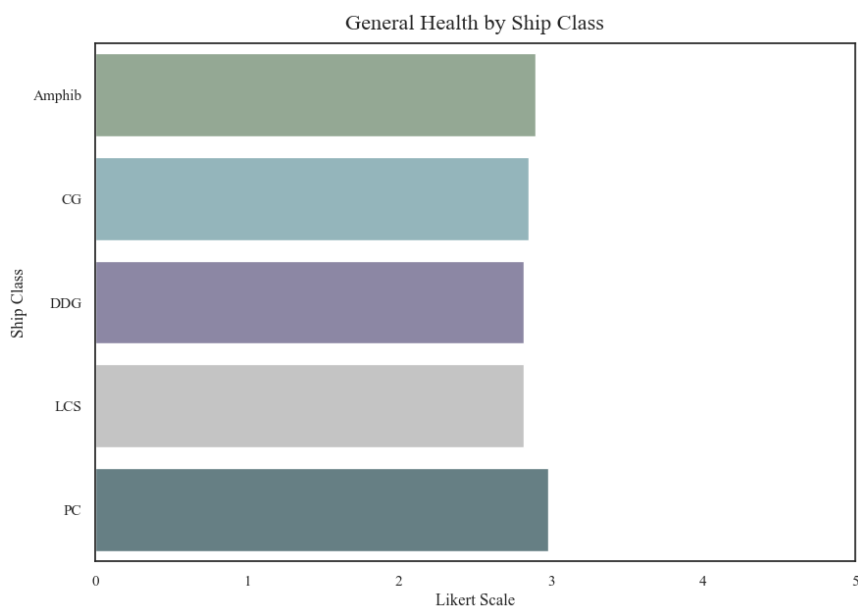


Figure 12: Bar Chart of General Health by Ship Class

General health was consistent across all ship classes. Overall, PC and Amphibs had less favorable responses toward general health

4.4.2.7 Mental Health

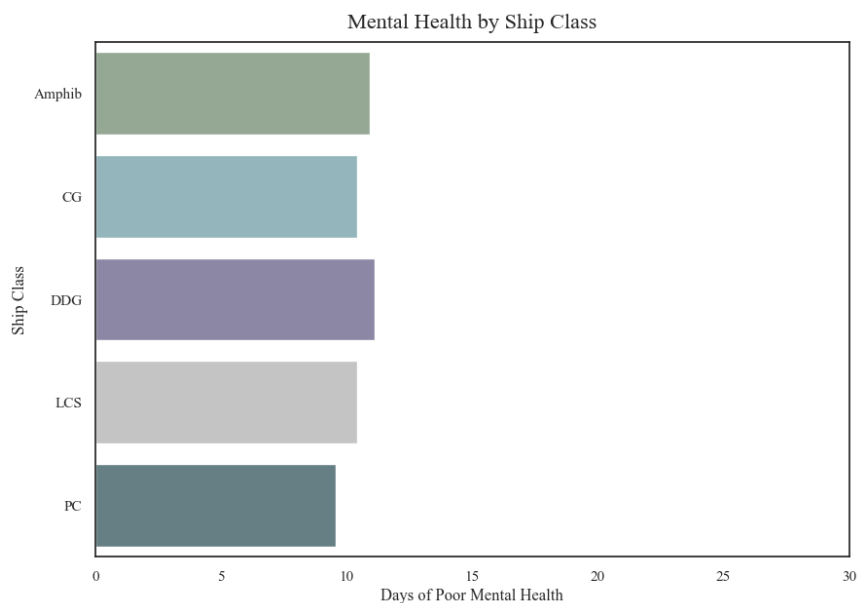


Figure 13: Bar Chart of Mental Health by Ship Class

Mental health was reported similarly across ship classes, but DDG reported more mental health concerns with Amphibs following.



4.4.2.8 Physical Health

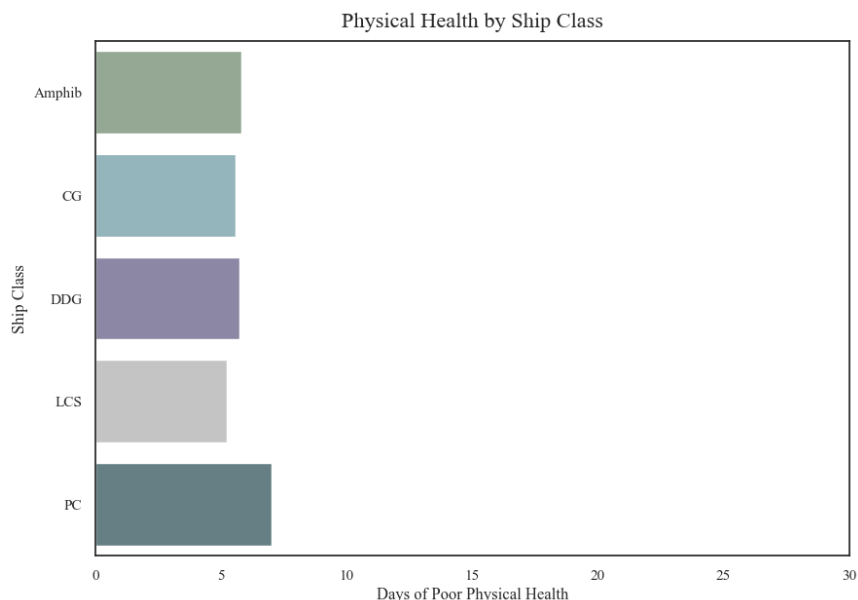


Figure 14: Bar Chart of Physical Health by Ship Class

Physical health was relatively consistent across ship classes. Unlike the other classes, PC reported more days of poor physical health with on average 7 days of poor physical health.

4.4.3 BREAKDOWN OF SLEEP FACTORS

The Sleep Factors construct is composed of 15 questions in the survey. This section examines the difference in sleep factors between ship classes. The factors with the largest negative impact for any individual group are shown to provide a better understanding of reasons for poor sleep. The Sleep Impairment and Sleep Disturbance Constructs are also shown for reference.



4.4.3.1 Sleep Factors of Highest Concern

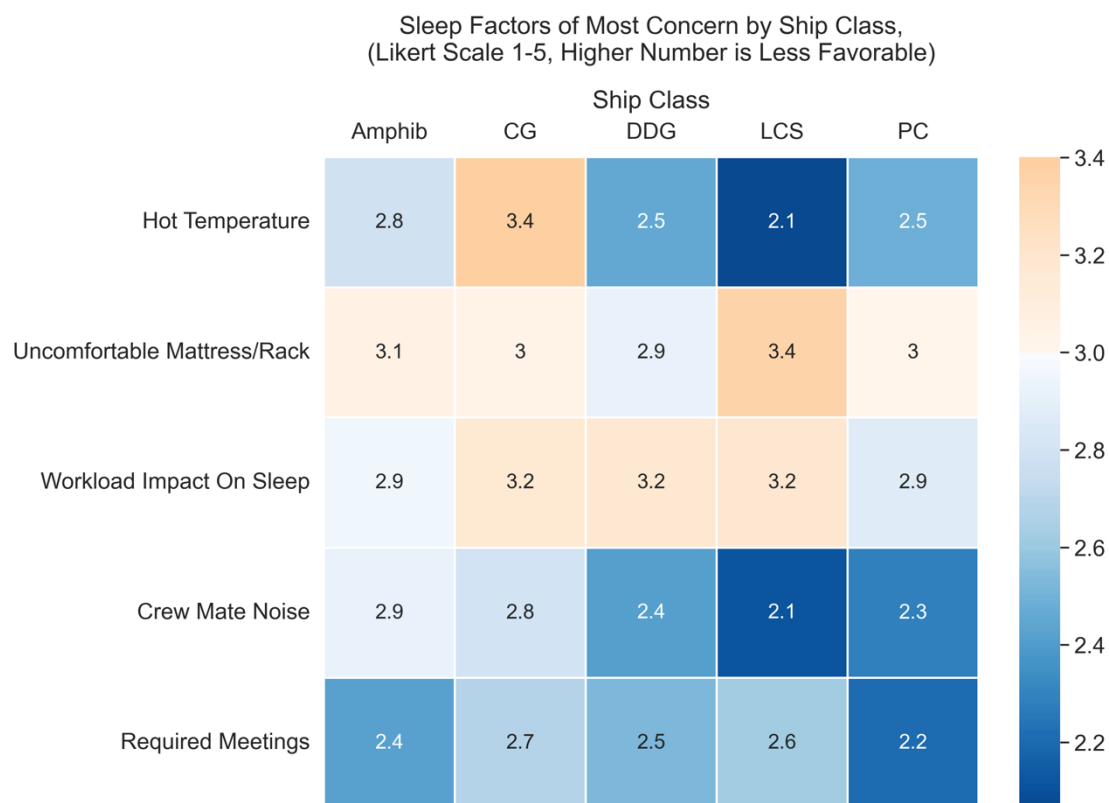


Figure 15: Heatmap of Sleep Factors by Ship Class

Hot Temperature for Cruisers was the most negative factor for any group, though the other ship classes reported less negatively for this factor. Uncomfortable Mattress/Rack was most negative for Littoral Combat Ships with other ship classes responding slightly more favorably. Workload Impact on Sleep, Crew Mate Noise, and Required Meetings were additional factors impacting sleep.

4.4.3.2 Means Constructs Relating to Sleep

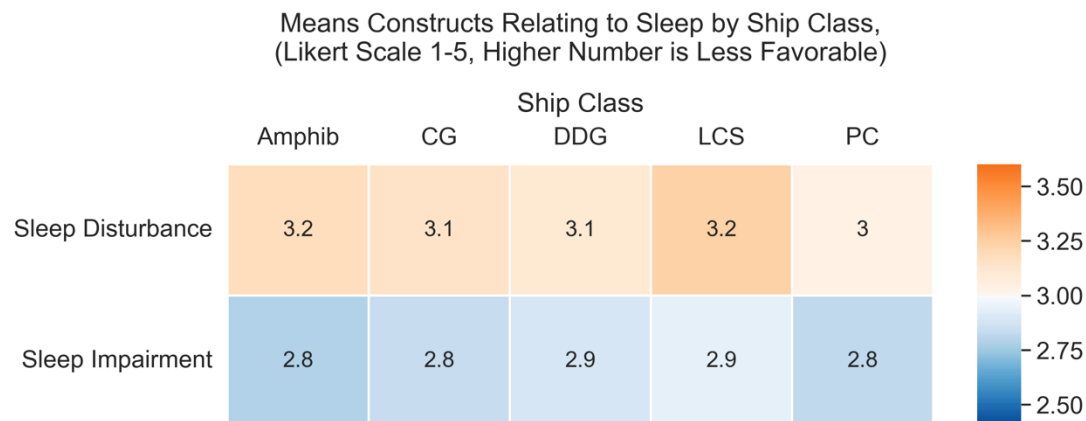


Figure 16: Heatmap of Sleep Constructs by Ship Class

Despite some pronounced differences in individual sleep factors, the Sleep Disturbance and Sleep Impairment constructs were relatively consistent across ship classes.



4.5 Paygrade Analysis

The following section examines the key constructs relative to rank: Junior Sailors (E1-E3), Petty Officers (E4-E6), Chiefs (E7-E9), and all Officers / Chief Warrant Officers (CWOs).

4.5.1 AREAS OF PRAISE AND CONCERN

Table 12: Constructs of Praise and Concern by Officer vs Enlisted

Item	Officer Mean	Non-Officer Mean
Leader Perception of Subordinates (Subset 3)	4.06	3.84
Unit Cohesion (Watch Team)	4.04	3.62
Personal Responsibility for Workplace Safety	3.98	3.49
Unit Cohesion (Work Center)	3.94	3.56
Supervisor Safety	3.83	3.53
Command Safety Practices	3.73	3.42
Leadership Perception of Subordinates (Subset 2)	3.71	3.71
Safety Satisfaction	3.69	3.27
Job Resources (Equipment)	3.69	3.38
Leadership Perception of Subordinates (Subset 1)	3.64	3.58
Job Stress *	3.70	3.62
Job Demands *	3.28	3.22
Affective Commitment	3.12	2.53
No-Blame Error Reporting	3.06	2.77
Team Process (Interpersonal Processes)	3.21	2.78
Satisfaction of Command	3.53	2.81
Psychological Safety	3.30	2.83
Team Process (Action Processes)	3.16	2.88
Team Processes (Overall)	3.21	2.88
Team Process (Transition Processes)	3.30	2.98

Note:

1. For the table above, higher scores represent a more favorable response, with a mean of 3 being neutral. Favorable scores are colored green, and unfavorable scores are shaded red. Scores within 0.05 of neutral (3) are shaded white.
2. *For these questions lower scores represent a more favorable response, with a mean of 3 being neutral.

4.5.2 CONSTRUCTS WITH NOTABLE DIFFERENCES IN MEANS

The means between the Officers and Enlisted Sailors differed for most constructs; the full results are provided in 6.2 Appendix B. The constructs below are those with the greatest differences between the mean scores. A test of statistical significance is given in Table 13 as well and all are significant with moderate effect sizes.



Table 13: Constructs with Significant Differences in Means Based on Officer vs Enlisted

Item	Officer			Enlisted			T-test		
	Mean	SD	Alpha	Mean	SD	Alpha	Mean Difference	P-Value	Effect Size
Unit Cohesion (Watch Team) Construct	4.04	0.77	0.93	3.62	0.94	0.94	0.42	<0.001***	0.489
Leader Assessment Construct	3.64	0.83	0.76	3.18	0.88	0.76	0.46	<0.001***	0.538
Affective Commitment Construct	3.12	1.09	0.91	2.53	1.09	0.91	0.59	<0.001***	0.541
Error Management Construct	3.66	0.82	0.69	3.26	0.89	0.70	0.4	<0.001***	0.467
Psychological Safety Construct	3.30	0.73	0.76	2.83	0.81	0.79	0.47	<0.001***	0.610
Safety Satisfaction Construct	3.69	0.76	0.88	3.27	0.89	0.90	0.42	<0.001***	0.508
Personal Responsibility for Workplace Safety Construct	3.98	0.76	0.87	3.49	0.86	0.86	0.49	<0.001***	0.604
Satisfaction with Command Construct	3.53	1.07	0.81	2.81	1.17	0.84	0.72	<0.001***	0.642
Days which mental health was judged 'not good' *	7.57	8.89		11.42	10.92		-3.85	<0.001***	0.387
Days which physical health was judged 'not good'*	4.02	7.07		5.98	8.79		-1.96	<0.001***	0.246

* For these items lower scores represent a more favorable response

The means above highlight the constructs with the most considerable mean differences between Officer and Enlisted Sailors. The constructs were all statistically significant between the groups and were of medium effect size except mental and physical health, which had a smaller effect size. Officers had more favorable responses towards every item.

4.5.3 CONSTRUCT DEEP DIVE

The aforementioned key constructs are analyzed by rank, with resolution beyond Officer and Enlisted groups. Table 14 details the construct means.

Table 14: Construct Means by Paygrade

Construct	Paygrade			
	Junior Sailors	Petty Officers	Chiefs	Officer/CWO
Sleep Disturbance *	3.23	3.47	3.37	2.98
Sleep Factors *^	2.49	2.55	2.20	2.07
Sleep Impairment *	2.78	2.96	2.70	2.64
Job Stress *	3.42	3.66	3.67	3.70
Job Demands *	3.09	3.25	3.25	3.28
General Health *^	2.81	2.73	2.93	3.27
Mental Health *+	10.50	12.06	9.36	7.57
Physical Health *+	5.61	6.00	6.33	4.02

Note*: For these questions lower scores represent a more favorable response

Note *+ : For these questions, the values represent the Number of Days. They are also not constructs; therefore, a Cronbach Alpha calculation would not be relevant.

Note *^ : This question is not a construct; therefore, a Cronbach Alpha calculation would not be relevant.

While construct means are very similar, the distributions show variability as denoted in Figures 13-18. Additional statistical measures comparing the groups are detailed in Appendix B.



4.5.3.1 Sleep Disturbance

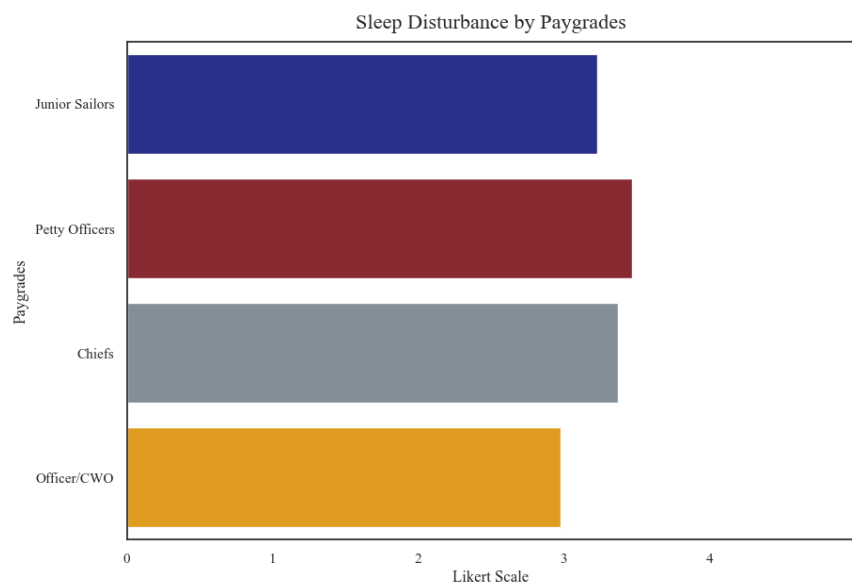


Figure 17: Bar Chart of Sleep Disturbance by Paygrade

Differences in Sleep Disturbance were reported statistically significant across different paygrades. Overall, Petty Officers and Chiefs reported more issues with Sleep Disturbance, compared to Officer/CWO reporting less issues.

4.5.3.2 Sleep Factors

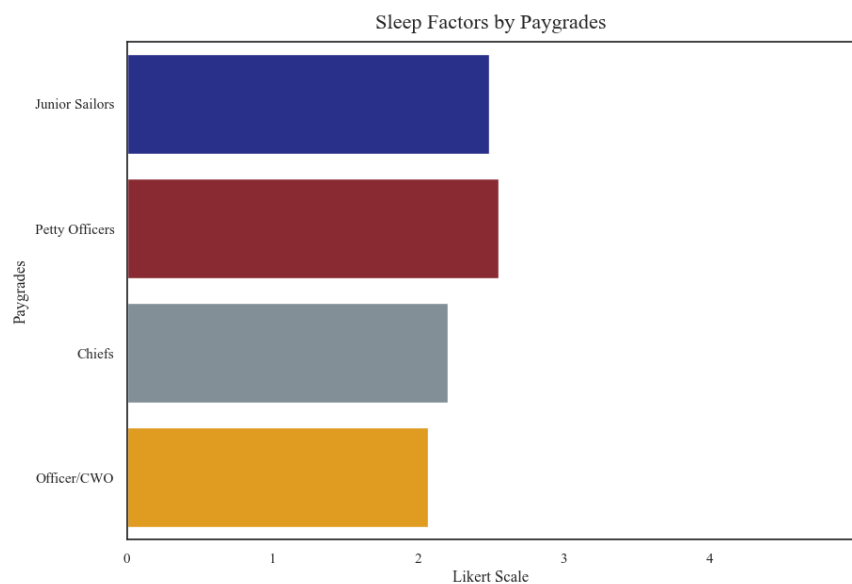


Figure 18: Bar Chart of Sleep Factors by Paygrade

Sleep factors had notable differences across paygrades. Petty Officers reported the most issues impacting Sleep Factors next to Junior Sailors.



4.5.3.3 Sleep Impairment

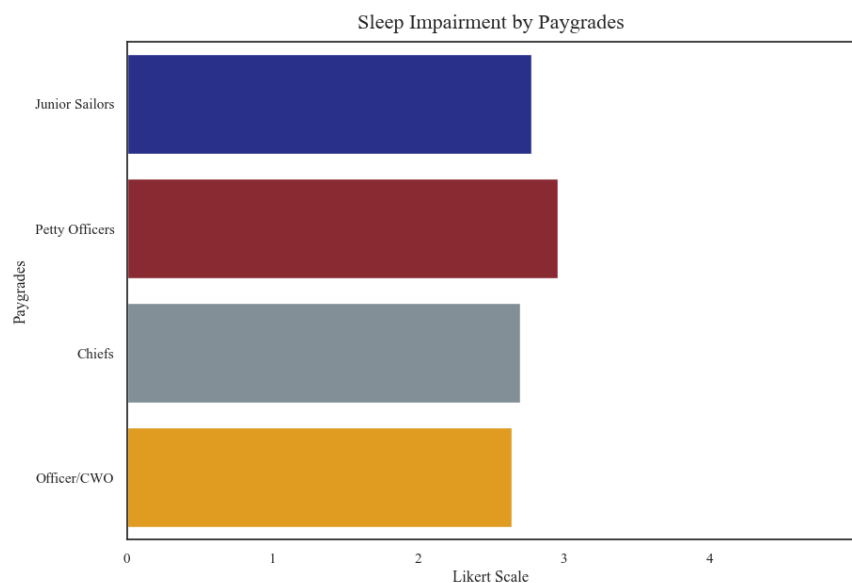


Figure 19: Bar Chart of Sleep Impairment by Paygrade

Sleep Impairment were reported similarly across paygrades. Petty Officers reported the most issues with sleep impairment compared to Officer/CWO with the least sleep impairment issues.

4.5.3.4 Job Stress

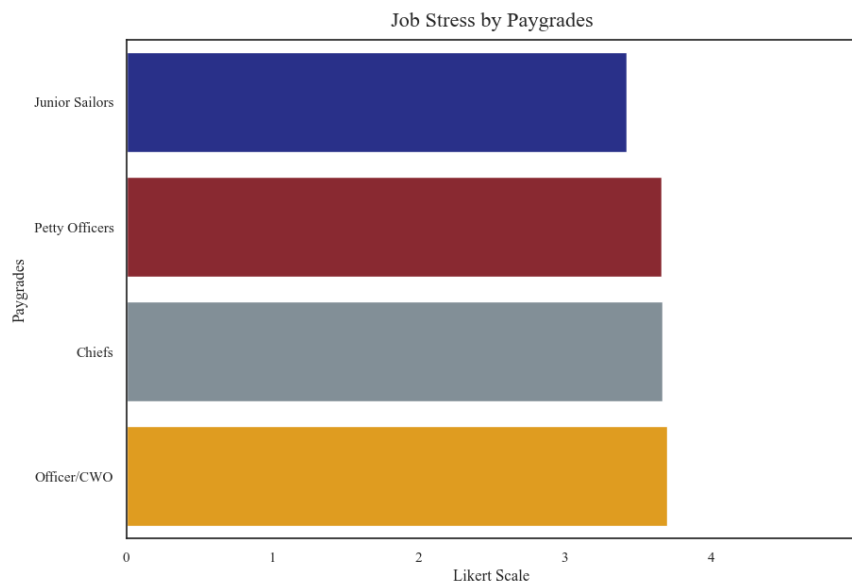


Figure 20: Bar Chart of Job Stress by Paygrade

Job Stress showed slight differences between pay grades. Junior Sailors on average reported less stress compared to Officer/CWO, Chiefs and Petty Officers.



4.5.3.5 Job Demand

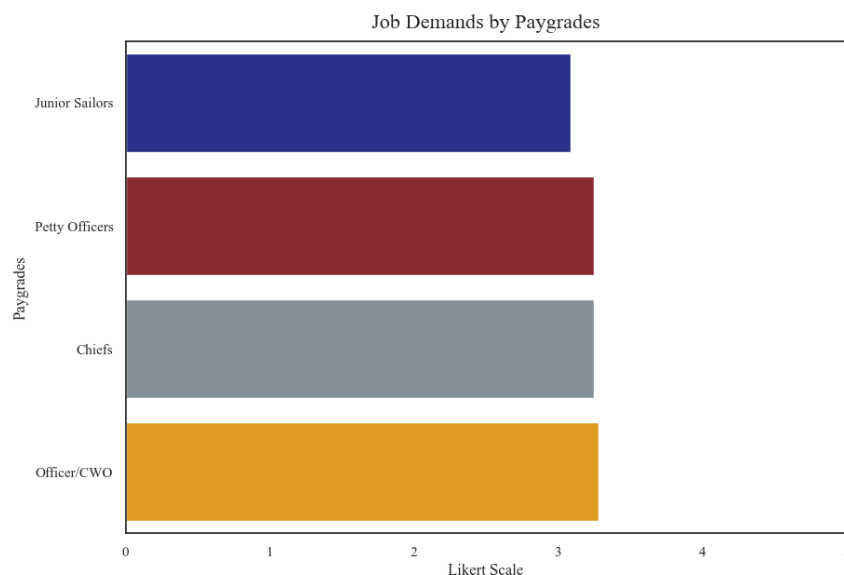


Figure 21: Bar Chart of Job Demands by Paygrade

Similar to Job Stress, Job Demand had the same trend across pay grades. Officer/CWO, Chiefs, and Petty Officers reported higher job stress compared to Junior Sailors.

4.5.3.6 General Health

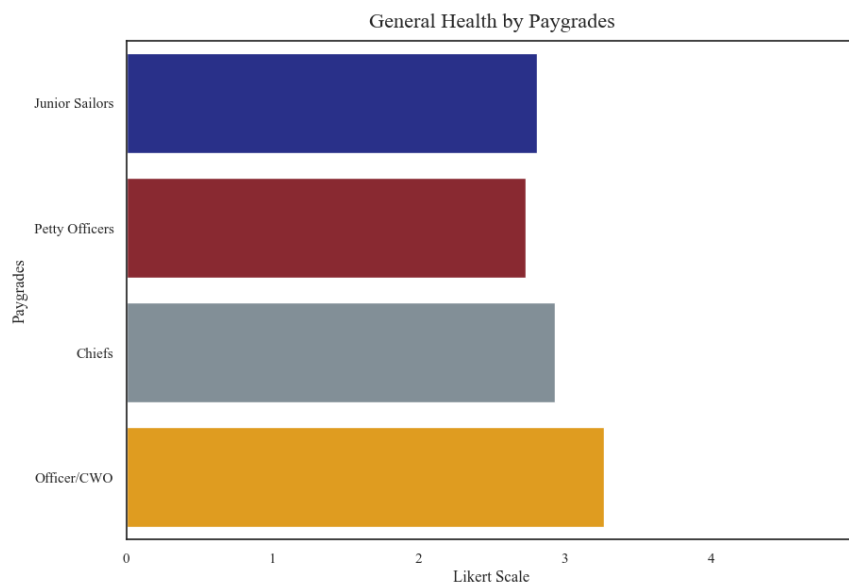


Figure 22: Bar Chart of General Health by Paygrade

General Health also had statistically significant differences across pay grades. Petty Officers reported slightly better general health next to Junior Sailors. Officer/CWO reported the worst general health overall.



4.5.3.7 Mental Health

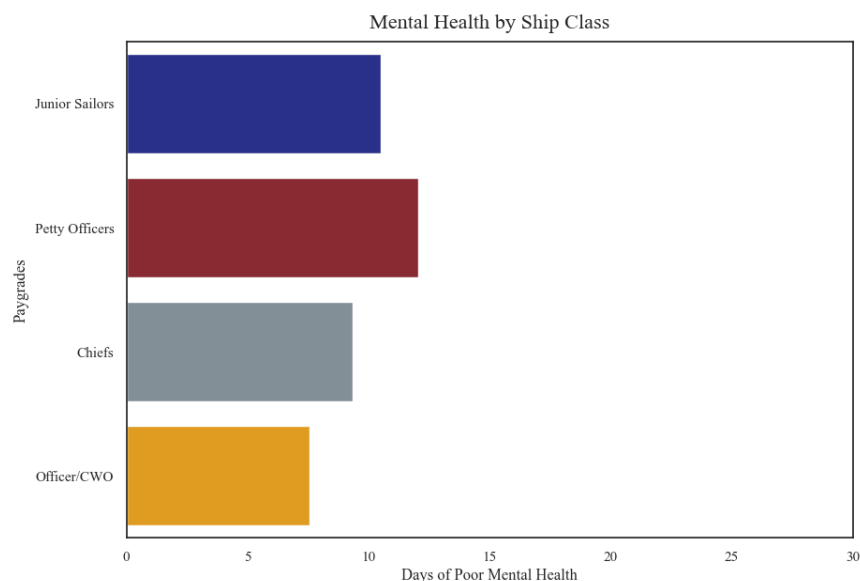


Figure 23: Bar Chart of Mental Health by Paygrade

Mental health was notably different across different pay grades. Petty Officers reported more poor days of mental health next to Junior Sailors.

4.5.3.8 Physical Health

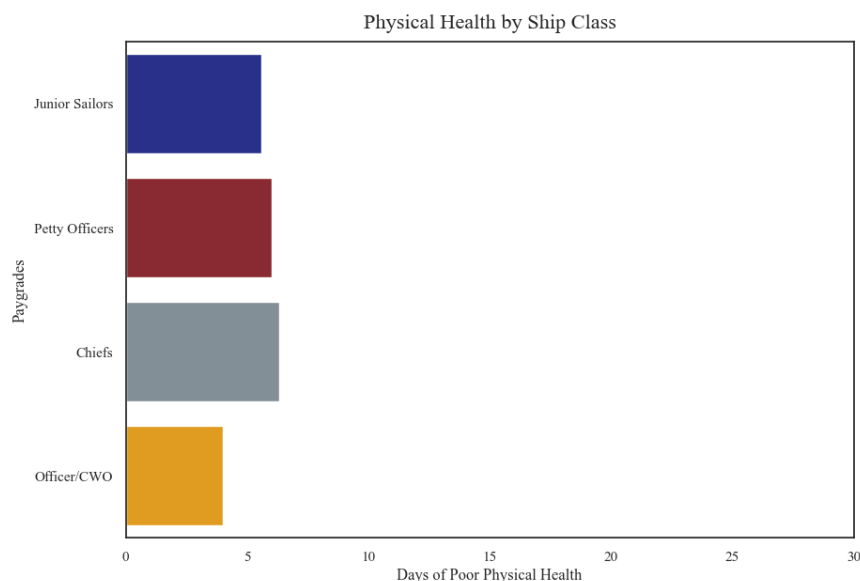


Figure 24: Bar Chart of Physical Health by Paygrade

Physical Health had notable differences across pay grades. Chiefs reported the most days of poor physical health, whereas Officer/CWO reported the least days of poor physical health.



4.5.4 BREAKDOWN OF SLEEP FACTORS

The Sleep Factors construct is composed of 15 questions in the survey. This section examines the difference in sleep factors between the Enlisted (E1-E9) group and the Officer group. The factors with the largest negative impact for either group are shown to provide a better understanding of reasons for poor sleep. The Sleep Impairment and Sleep Disturbance Constructs are also shown for reference.

4.5.4.1 Sleep Factors of Highest Concern

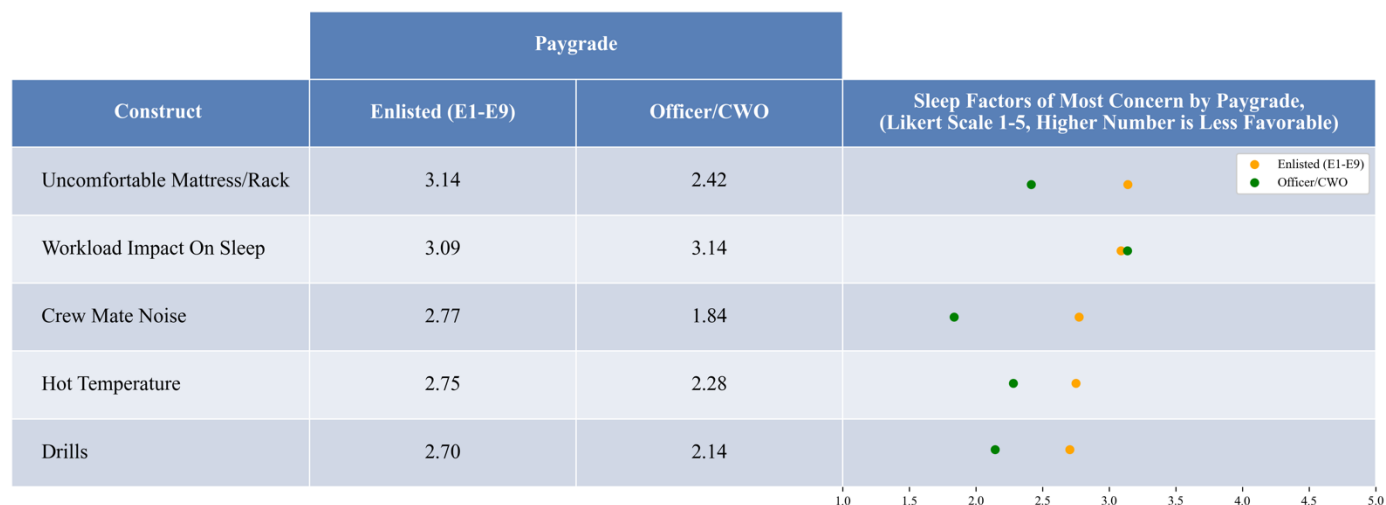


Figure 25: DNA Chart of Sleep Factors by Paygrade

Uncomfortable Mattress/Rack was the largest negative factor for Enlisted crewmembers and showed a pronounced difference between the Enlisted and Officer group. Workload was a large negative factor but showed a negligible difference between groups. Crew Mate Noise, Hot Temperature, and Drills all were highly negative and showed a large difference between groups, with Enlisted consistently reporting more negative results.

4.5.4.2 Sleep Constructs of Highest Concern

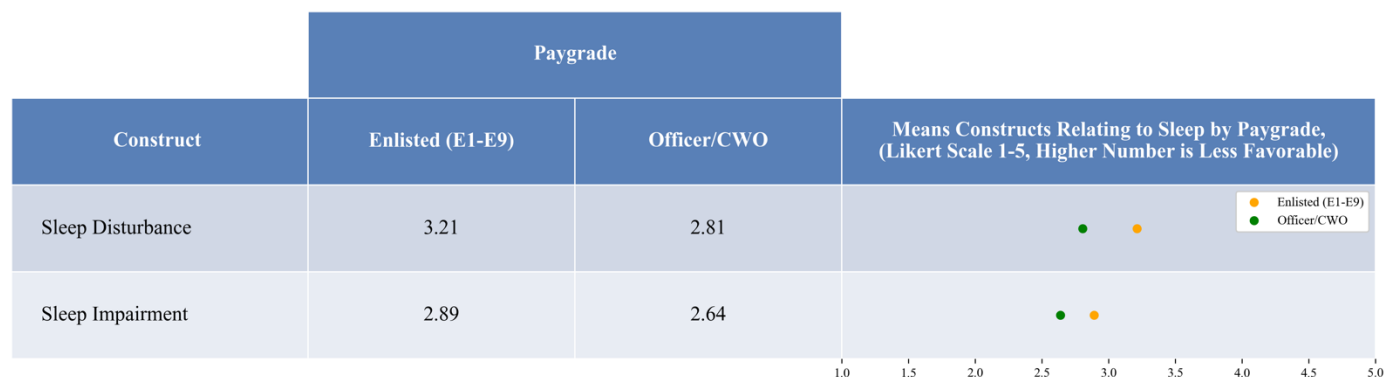


Figure 26: DNA Chart of Sleep Constructs by Paygrade

Consistent with the highlighted sleep factors, the Sleep Disturbance and Sleep Impairment constructs showed differences between groups, with the Enlisted group reporting less favorable results than the Officer group.



4.5.5 ACTIVITY ANALYSIS

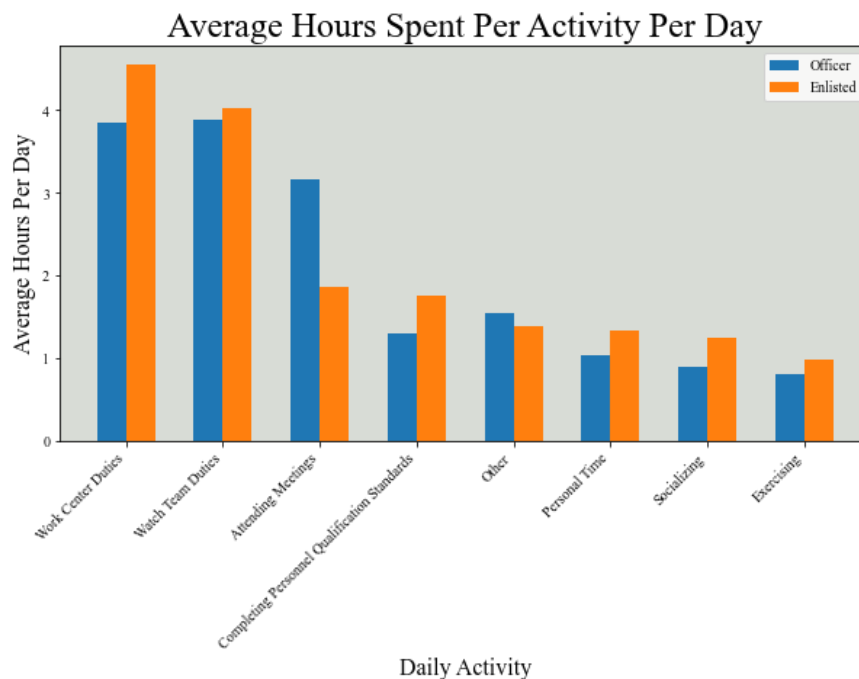


Figure 27: Average Hours of Spent per Activity per Day (Paygrade)

Officers/CWOs reported spending an average of 12 hours and 11 minutes on work-related activities and 4 hours and 15 minutes on non-work-related activities. Enlisted Sailors reported spending 12 hours and 11 minutes on work-related activities and 4 hours and 56 minutes on non-work-related activities.

4.6 Department Analysis

Analysis was performed to identify differences in responses among the various departments. All results include responses from 12 departments (Engineering, Operations, Combat Systems, Supply, Admin, Deck, Weapons, Medical/Dental, Training / Plans & Tactics, Air, AIMD and, Safety). It was noted that Engineering department responses were significantly different from all other departments. Additionally, many of the departments had too few responses to allow for a robust comparative analysis. As such, departmental analysis was conducted by comparing all ship departments to the Engineering department.



4.6.1 AREAS OF PRAISE BY DEPARTMENT

Table 15: Ship Department Areas of Praise and Concern

Item	Engineering Mean	Non-Engineering Mean
Leader Perception of Subordinates (Subset 3)	3.77	3.93
Unit Cohesion (Watch Team)	3.69	3.69
Unit Cohesion (Work Center)	3.66	3.61
Leadership Perception of Subordinates (Subset 2)	3.63	3.73
Personal Responsibility for Workplace Safety	3.58	3.57
Leadership Perception of Subordinates (Subset 1)	3.54	3.61
Supervisor Safety	3.51	3.60
Command Safety Practices	3.36	3.50
Job Resources (Equipment)	3.34	3.46
Affective Commitment	2.58	2.64
Job Stress *	3.81	3.58
Job Demands *	3.42	3.17
Team Process (Interpersonal Processes)	2.75	2.87
No-Blame Error Reporting	2.81	2.82
Team Process (Action Processes)	2.84	2.95
Team Processes (Overall)	2.85	2.96
Satisfaction of Command	2.87	2.95
Psychological Safety	2.88	2.92

Note:

1. For the table above, higher scores represent a more favorable response, with a mean of 3 being neutral. Favorable scores are colored green, and unfavorable scores are shaded red. Scores within 0.05 of neutral (3) are shaded white.

2. *For these questions lower scores represent a more favorable response, with a mean of 3 being neutral.

4.6.2 CONSTRUCTS WITH NOTICEABLE DIFFERENCES IN MEANS

Compared across different constructs, notable differences were not observed between Engineering and Non-Engineering departments. Engineering reported slightly worse physical and mental health days than Non-Engineering departments.

Table 16: Major Mean Differences between Engineering and Non-Engineering Departments

Item	Engineering			Non-Engineering			T-test		
	Mean	SD	Alpha	Mean	SD	Alpha	Mean Difference	P-Value	Effect Size
Job Stress Construct *	3.81	0.90	0.84	3.58	0.95	0.84	0.23	<0.001***	0.249
Occupational Challenges Construct *	2.91	1.00	0.73	2.70	0.99	0.72	0.21	0.001**	0.211
Safety Satisfaction Construct	3.22	0.89	0.90	3.37	0.88	0.90	-0.15	<0.001***	0.169
Job Demands Construct *	3.42	0.90	0.84	3.17	0.88	0.83	0.25	<0.001***	0.281
Sleep Impairment Construct *	3.02	0.98	0.90	2.80	0.92	0.89	0.22	<0.001***	0.231
Days which mental health was judged 'not good' *	11.80	11.05		10.62	10.65		1.18	0.132	0.109
Days which physical health was judged 'not good'*	6.89	9.32		5.37	8.32		1.52	0.005*	0.172

* For these items lower scores represent a more favorable response

The above table reports the constructs with the most considerable mean differences between the Engineering and other departments. The constructs were all statistically significant between the groups but had a small effect size. The non-Engineering department had favorable responses towards every item.

4.6.3 CONSTRUCT DEEP DIVE

This analysis compares means between our constructs of interests and the ship departments, which were dichotomized between engineering and non-engineering departments. All differences were statistically significant except the Sleep Disturbance construct.

Table 17: Constructs Means and Tests of Significance based on Ship Departments.

Ship Department (Engineering vs. Non-Engineering)			
Construct	Mean Difference	P-Value	Effect Size
Sleep Disturbance *	-0.01	0.61	0.01
Sleep Factors *^	-0.08	<0.01	0.06
Sleep Impairment *	-0.21	<0.01	0.20
Job Stress *	-0.22	<0.01	0.21
Job Demands *	-0.25	<0.01	0.21
General Health *^	-0.15	<0.01	0.18
Mental Health *+	-1.18	0.01	0.11
Physical Health *+	1.52	<0.01	-0.16

Note*: For these questions lower scores represent a more favorable response

Note *+ : For these questions, the values represent the Number of Days. They are also not constructs; therefore, a Cronbach Alpha calculation would not be relevant.

Note *^ : This question is not a construct; therefore, a Cronbach Alpha calculation would not be relevant.

4.6.4 BREAKDOWN OF SLEEP FACTORS

This section examines the difference in sleep factors between Engineering and Non-Engineering departments. The factors with the largest negative impact for either group are shown to provide a better understanding of reasons for poor sleep. The Sleep Impairment and Sleep Disturbance Constructs are also shown for reference.

4.6.4.1 Sleep Factors of Highest Concern

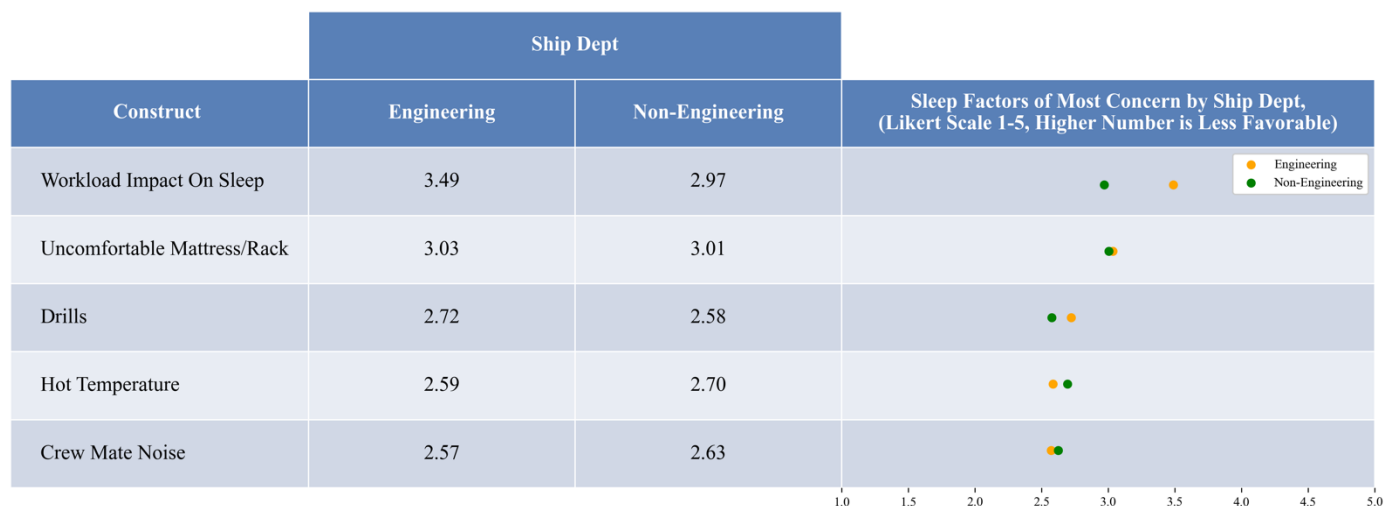


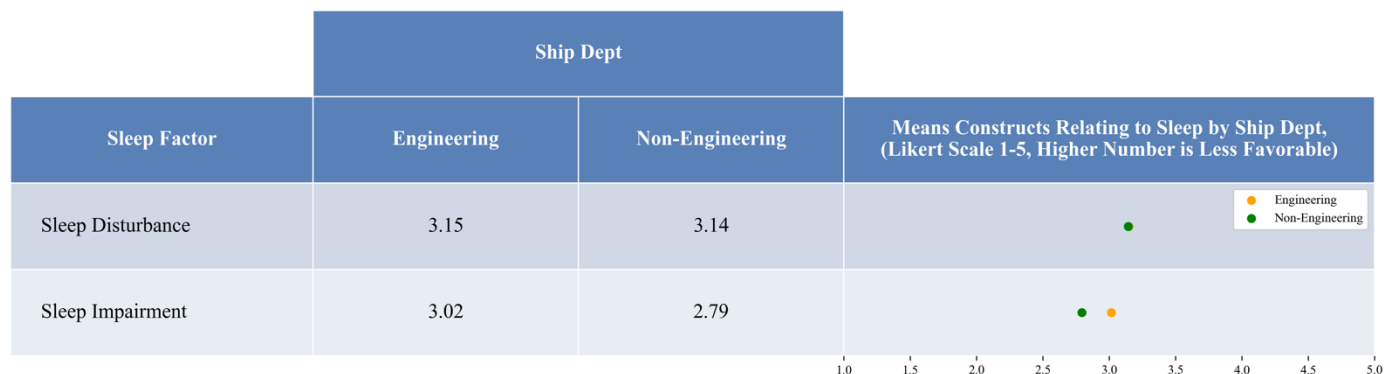
Figure 28: Sleep Factors by Ship Department

Workload Impact on Sleep was the largest negative factor for the Engineering Department and showed a pronounced difference between the Engineering and Non-Engineering departments.



Uncomfortable Mattress/Rack, Drills, Hot Temperature, and Crew Mate Noise were additional factors impacting sleep, but differences between groups were minor.

4.6.4.2 Sleep Constructs of Highest Concern



The Sleep Impairment construct showed a difference between groups, with the Engineering department reporting less favorable results than the Non-Engineering departments. The Sleep Disturbance construct showed a negligible difference between groups.

4.6.5 ACTIVITY ANALYSIS

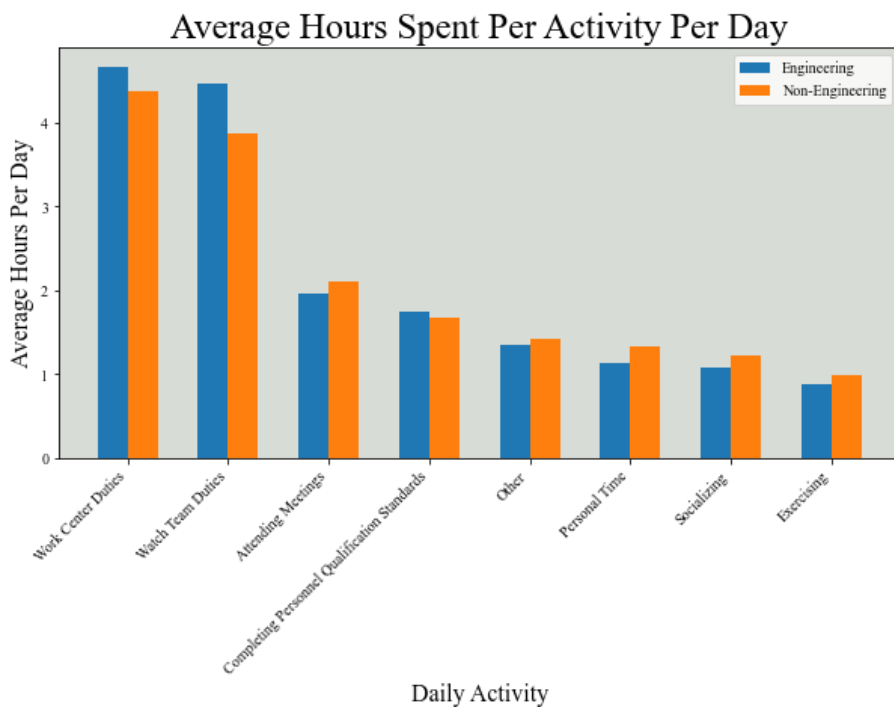


Figure 30: Average Hours Spent per Activity per Day (Department)

Those in Engineering departments reported spending an average 12 hours and 55 minutes on work-related activities and 4 hours and 23 minutes on non-work-related activities. All other departments reported spending 12 hours and 3 minutes on work-related activities and 4 hours and 56 minutes on non-work-related activities. Engineering department personnel reported more time being focused on work center and Watch Team duties and less time attending meetings; as well as less time per day devoted to recuperative activities (e.g., personal time, socializing and exercising).



4.7 Fit/Fill Analysis

4.7.1 SUMMARY

FIT and FILL are a good criterion variable and choice of threshold will separate a sample into two groups, a *top* and *bottom* group, with the top group exhibiting more desirable characteristics than the other (e.g., less work stress, better sleep, etc.). The amount of required staffing in an operational environment links to certain variables (e.g., a strong association with work demands and leadership assessments). This analysis explored this variation, across several variables that potentially connect to manning requirements. FIT is found to be a better criterion variable than FILL as it produces more pronounced differences in the top and bottom groups, often in the expected direction (e.g., the top group experiences less work stress than the bottom group). The choices of threshold (0.92 for FIT, 0.95 for FILL) are good, often producing the most pronounced separation in the two groups, in the ideal direction, though this claim holds most strongly for the FIT criterion variable. Detailed information is provided on the sample sizes that such cuts create and the corresponding measures of magnitude and statistical reliability on the outcome variables of interest (as provided by Cohen's *d* and tests of statistical significance), which can provide more nuanced guidance as how to set "acceptable" manning thresholds.

In summary, the FIT/FILL analysis suggests that ships with lower FIT metrics (less than 92%) reported less favorable responses than those with metrics above 92%. In addition, ships with lower FILL metrics (less than 95%), or low manning showed an increase in operational risk. Initial FIT/FILL analysis did not show a notable relationship between FIT/FILL and the constructs but further analysis is ongoing to define the boundaries of risk associated with manpower and manning mismatches.

4.7.2 EXPLORATORY ANALYSIS

This section explores the FIT/FILL criteria, and whether existing standards (0.92 for FIT, 0.95 for FILL) are accomplishing the goals they are intended to realize. Broadly speaking, these thresholds are intended to identify two subgroups, one with greater "preparedness" or readiness to meet mission demands than the other. There are different ways to assess "readiness", not all of which are expected to separate the two groups equally. For example, one might expect that separating ships into those that meet the criterion for sufficient, qualified manning would lead to two groups that differ in perceived work demands. Alternatively, one might expect that affective commitment to the organization might be influenced by adequate manning, but the link is less direct, and more tenuous than the link to work pressure.

What are criterion thresholds intended to accomplish? The three possibilities considered in this section are:

- (1) The group that meets the criterion should perform better or have more of the desired quality than those groups that fail to meet the criterion. (Direction of the difference)
- (2) This difference between the groups should be statistically reliable, and not simply due to chance variation. (Statistical robustness)
- (3) Since different thresholds create different subgroups, one would expect the magnitude of the difference to change across possible thresholds. (Variation in differences)



These three possibilities imply that a satisfactory criterion choice will produce a large separation in the two groups, in the desired direction of greater operational readiness. The following method will explore these properties across alternative choices of criterion (to include 0.92 for FIT and 0.95 for FILL, the currently enacted thresholds):

- (1) Select a range of possible criterion choices that include the existing ones as special cases: for FIT, consider 0.80 to 0.95, where 0.92 is the currently enacted threshold; for FILL, consider 0.80 to 0.99, where 0.95 is the current threshold. The analysis will test the foregoing hypotheses across each of these levels.
- (2) Identify a relevant construct that is hypothesized to be related to manning levels (e.g., work stress). If there are multiple items that capture some aspect of the same latent construct, create a composite score in which larger values mean that more of that property is present in a given unit (e.g., a ship or individual Sailor). The constructs considered in this section include:
 - a. Job stress
 - b. Job demands
 - c. Sleep obtained
 - d. Sleep deficiency
 - e. Workload as a factor affecting sleep
 - f. Fatigue-related impairment
 - g. Sleep disturbance
 - h. Refreshing sleep obtained
 - i. Time available to exercise
 - j. Leadership assessment
 - k. Affective commitment
 - l. Command satisfaction
 - m. Command safety
 - n. Social support
- (3) Apply a candidate criterion threshold to the data, which will generate a separation of the data into two subgroups: those units that meet or exceed the threshold, and those that do not. Calculate the arithmetic mean of each group, calculate their difference, and test whether this difference is in the appropriate direction, reflecting more of the desired property in the group that satisfies the criterion. A good choice for criterion threshold will produce the appropriate direction of the mean difference.
- (4) Calculate the statistical robustness of the difference using an independent samples t-test (not assuming equal variances, which is a more conservative test), and take the resulting p-value -- if it is less than 0.05 -- as evidence that the difference is statistically robust. A good choice for criterion threshold will capture stable differences between the two subgroups, not transient variation due to random noise.
- (5) Calculate a measure of the size of the difference using a standardized mean difference (Cohen's *d*, which is a more appropriate quantity than a raw difference in means, as it accounts for the inherent variability in the two subgroups). This value can be either negative or positive, depending on the direction of the difference in means, and absolute magnitudes of 0.20, 0.50, and 0.80 are often adopted as benchmarks for small, medium, and large differences. Plot these values across all choices of criterion values; compare the degree to which the two types of criteria -- FIT and FILL -- can produce distinguishable groups with the top group expressing greater "readiness" than the bottom group (as measured by the properties listed above). Note that values close to 0 indicate that the top and bottom groups are indistinguishable with respect to a particular property, at that given criterion value (e.g., job stress, command satisfaction, etc.). A good choice of



criterion will produce readily distinguishable groups, with the top group “outperforming” the bottom group, which includes having more of a desirable property, and less of an undesirable property.

IMPORTANT: Rules for interpreting each of the variables described in (2) will be provided in each section below, and some are repeated to help reduce the need to return to this section. In all cases, the absolute value of Cohen’s d measures the degree of separation between the top and bottom groups, but the desired direction (positive or negative Cohen’s d) can differ. That is, in some cases, a large positive Cohen’s d could indicate that the top group is outperforming the bottom group; in other cases, smaller values of Cohen’s d (e.g., $d = -0.81$) may reflect the same relationship.

CAVEAT: This is an exploratory analysis that uses tests of statistical significance as a means for addressing the question of “statistical robustness”, or the extent to which a difference may be judged real and not due simply to chance. However, relatively small sample sizes in some groups may undermine the ability of a test to detect a difference, even if it does exist, so results that fail to rule out null hypotheses of no difference should be considered as preliminary, provisional, and subject to change as additional data come in.

4.7.2.1 Job Stress

Job stress is a composite score—a sum across 5 items—with smaller values indicating less stress. Manning levels, if they are inadequate, may force greater work burdens on personnel, which could plausibly increase their job stress. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group exhibiting less job stress than the bottom group.

Figure 21 displays Cohen’s d , with negative values indicating that the subgroup satisfying the threshold exhibits less of the undesired property (i.e., less stress). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to job stress (Cohen’s $d \approx 0$). Further, values of Cohen’s $d < 0$ also indicate the most plausible direction: that the top group is expressing, on average, less stress than the bottom group.

For the FIT criterion, all differences are in the expected direction (all Cohen’s d values < 0), but only FIT thresholds at 0.90 and above are statistically significant ($p < 0.05$). Of these values, the threshold of 0.93 exhibited the smallest difference (-0.91), with 0.91 and 0.92 (the currently enacted threshold) close behind (-0.89 and -0.75, respectively). For the FILL criterion, all but two of the thresholds (0.80 and 0.81) are in the expected direction, but none of the differences displayed are statistically significant. and at their respective thresholds, both groups are essentially indistinguishable. However, for FILL at 0.95 (the current threshold), there is a suggestive, nontrivial difference ($d = -0.30$), but for now, there is insufficient evidence to suggest that it is stable one ($p > .05$).

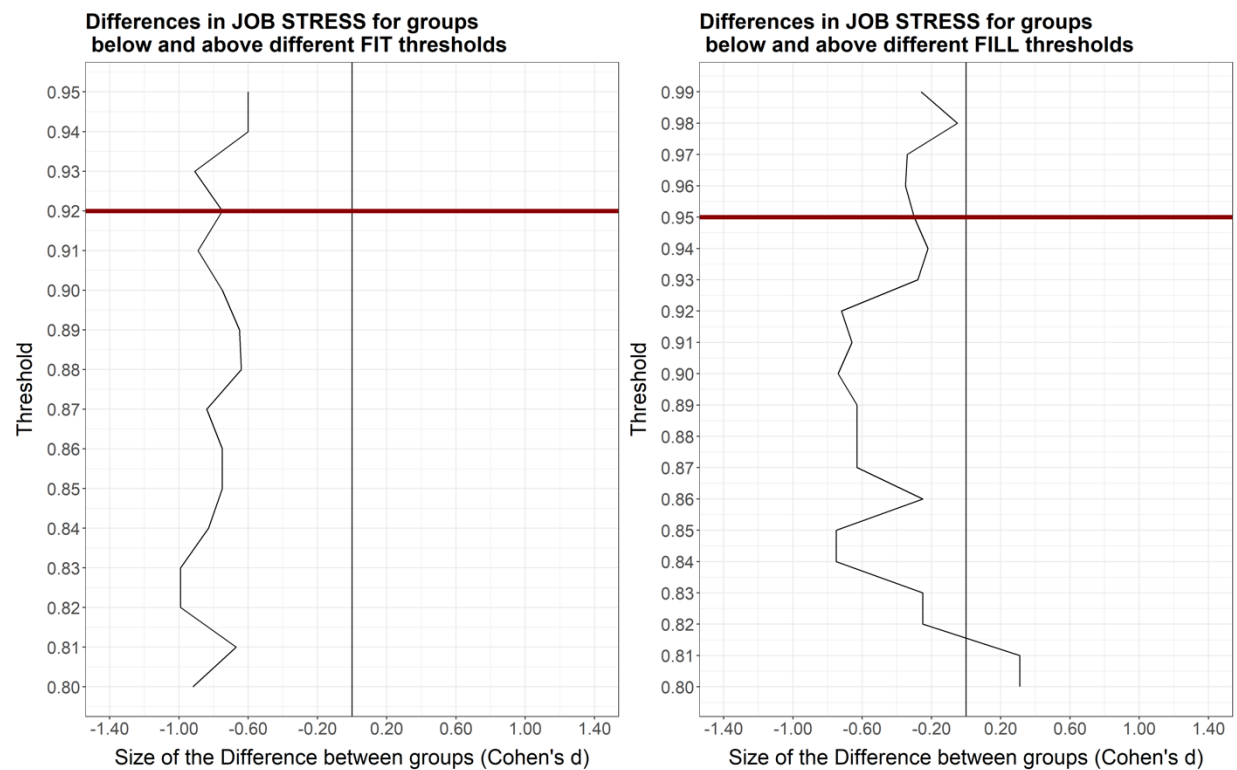


Figure 31 Comparison of Job Stress Across Different Fit and Fill Criterion Thresholds.

4.7.2.2 Job Demands

“Job demands” is a composite score -- a sum across 5 items -- with smaller values indicating fewer demands and less job burden. Manning levels, if they are inadequate, may force more job responsibilities on personnel, which could plausibly increase job demands. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group implicating workload less often than the bottom group as a cause of sleep issues.

Figure 22 displays Cohen’s d , with negative values indicating that the subgroup satisfying the threshold exhibits less of the undesirable property (i.e., burdensome job demands). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to job demands (Cohen’s $d \approx 0$). Further, values of Cohen’s $d < 0$ also indicate the most plausible direction.

For the FIT criterion, all differences are in the expected direction (all Cohen’s d values < 0), but only values between 0.89 and 0.93 are statistically significant ($p < 0.05$), with the thresholds of 0.87 and 0.88 both approaching statistical significance ($p = 0.065$ and $p = 0.057$, respectively). Of these values, the threshold of 0.91 exhibited the smallest difference (-1.11), with thresholds of 0.89, 0.90, 0.92, and 0.93 producing large differences in the two groups (-0.81, -0.80, -0.84, -0.87, respectively). For the FILL criterion, all but two thresholds (0.80 and 0.81) are in the expected direction, but none of the differences displayed are statistically significant. However, for FILL at 0.95 (the current threshold), there is a suggestive, nontrivial difference ($d = -0.25$), but for now, there is insufficient evidence to suggest that it is stable one ($p > .05$).

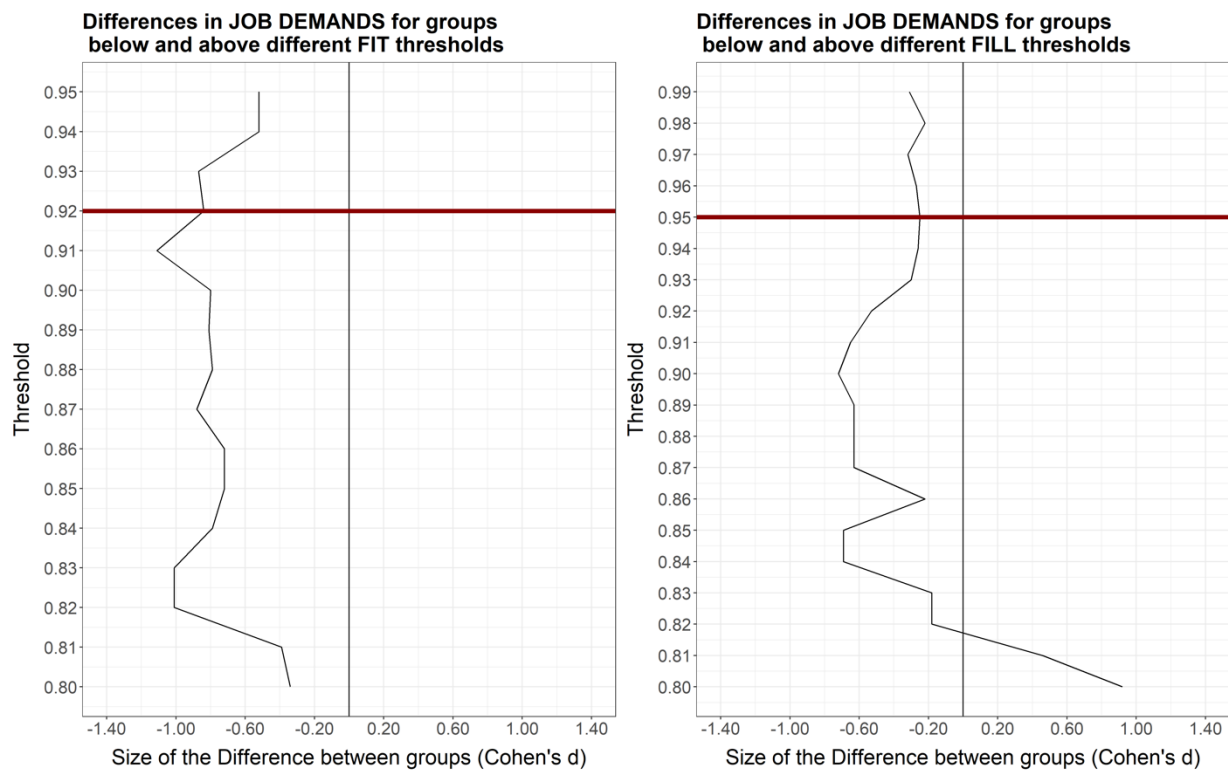


Figure 32 Comparison of Job Demands Across Different Fit and Fill Criterion Thresholds



4.7.2.3 Sleep Obtained

“Sleep obtained” refers to personnel estimates of the amount of sleep they obtain onboard a ship. Manning levels, if they are inadequate, may force more job responsibilities on personnel, which could plausibly reduce the amount of sleep obtained. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group obtaining more sleep than the bottom group.

Figure 23 displays Cohen’s d , with positive values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., obtaining more sleep). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to sleep obtained (Cohen’s $d \approx 0$). Further, values of Cohen’s $d > 0$ indicate the most plausible direction.

For the FIT criterion, all differences are in the expected direction (all Cohen’s d values > 0), but only values from 0.92 to 0.95 approach or reach statistical significance ($p < 0.07$). Of these values, the threshold of 0.93 exhibited the largest difference (0.63), with the threshold of 0.92 producing a medium-sized difference (0.59). For the FILL criterion, all but one of the thresholds (0.81) are in the expected direction, but none of the differences displayed are statistically significant. However, for FILL at 0.95 (the current threshold), there is a suggestive, nontrivial difference ($d = 0.22$), and for FILL at 0.92, there is a potentially large difference ($d = 0.72$, $p = 0.10$). For now, however, there is insufficient evidence to suggest that either difference is stable ($p > .05$).

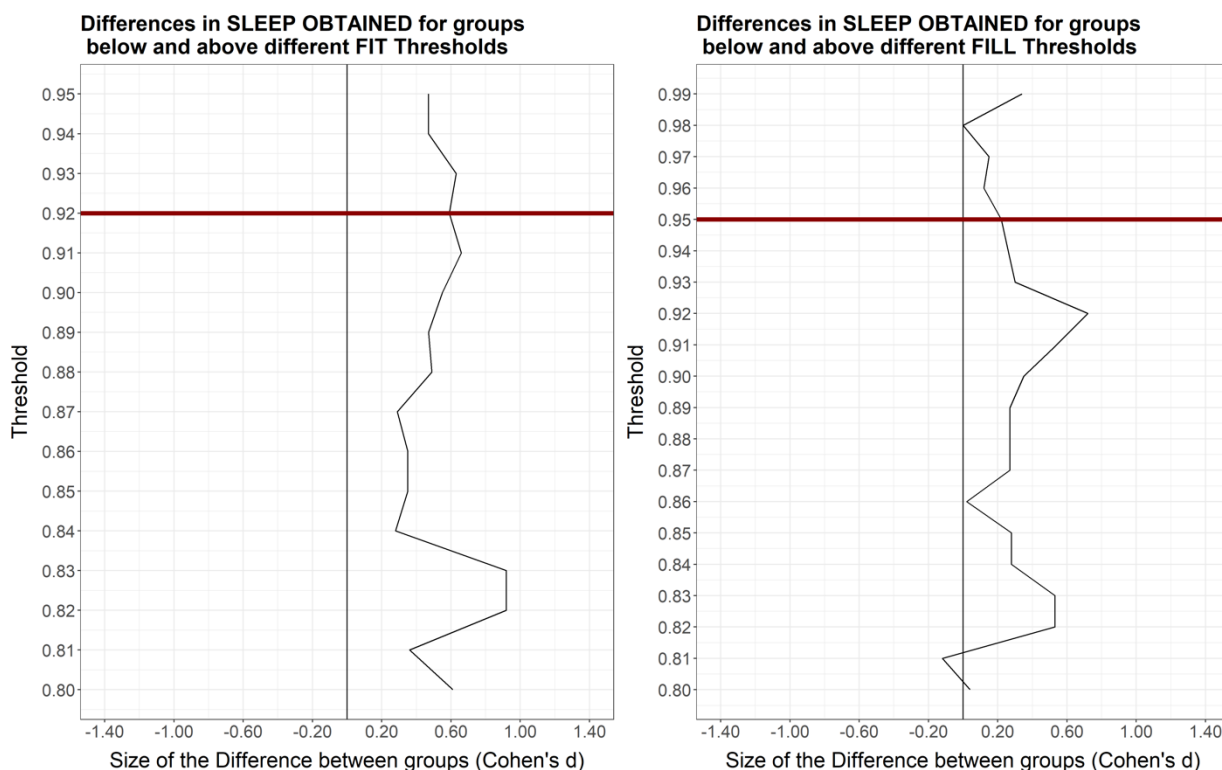


Figure 33 Comparison of Sleep Obtained Across Different Fit and Fill Criterion Thresholds



4.7.2.4 Sleep Deficiency

“Sleep deficiency” is a derived score -- a ratio of sleep obtained to sleep required -- that is intended to capture individual variation in the amount of sleep specific Sailors may need to function at the highest levels. Values less than 1 indicate that one is getting less sleep than one needs. Manning levels, if they are inadequate, may force greater work burdens on personnel, which could have an effect on their sleep. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group obtaining more sleep (adjusted to reflect individual needs) than the bottom group.

Figure 24 displays Cohen’s d , with positive values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., obtaining more sleep relative to one’s needs). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to sleep obtained (Cohen’s $d \approx 0$). Further, values of Cohen’s $d > 0$ indicate the most plausible direction.

For the FIT criterion, all differences are in the expected direction (all Cohen’s d values > 0), with thresholds of 0.90, 0.92, 0.94, 0.95 reaching statistical significance ($p < 0.05$). Of these values, the threshold of 0.90 produces the largest difference (0.71), with the threshold of 0.92 producing a medium-sized difference (0.60). For the FILL criterion, all but four of the thresholds (0.81, 0.86, 0.97, 0.98) are in the expected direction, but none of the differences displayed are statistically significant. However, for FILL at 0.95 (the current threshold), there is a suggestive, nontrivial difference ($d = 0.21$), and for FILL at 0.92, there is a potentially large difference ($d = 0.72$, $p = 0.12$). For now, however, there is insufficient evidence to suggest that either difference is stable ($p > .05$).

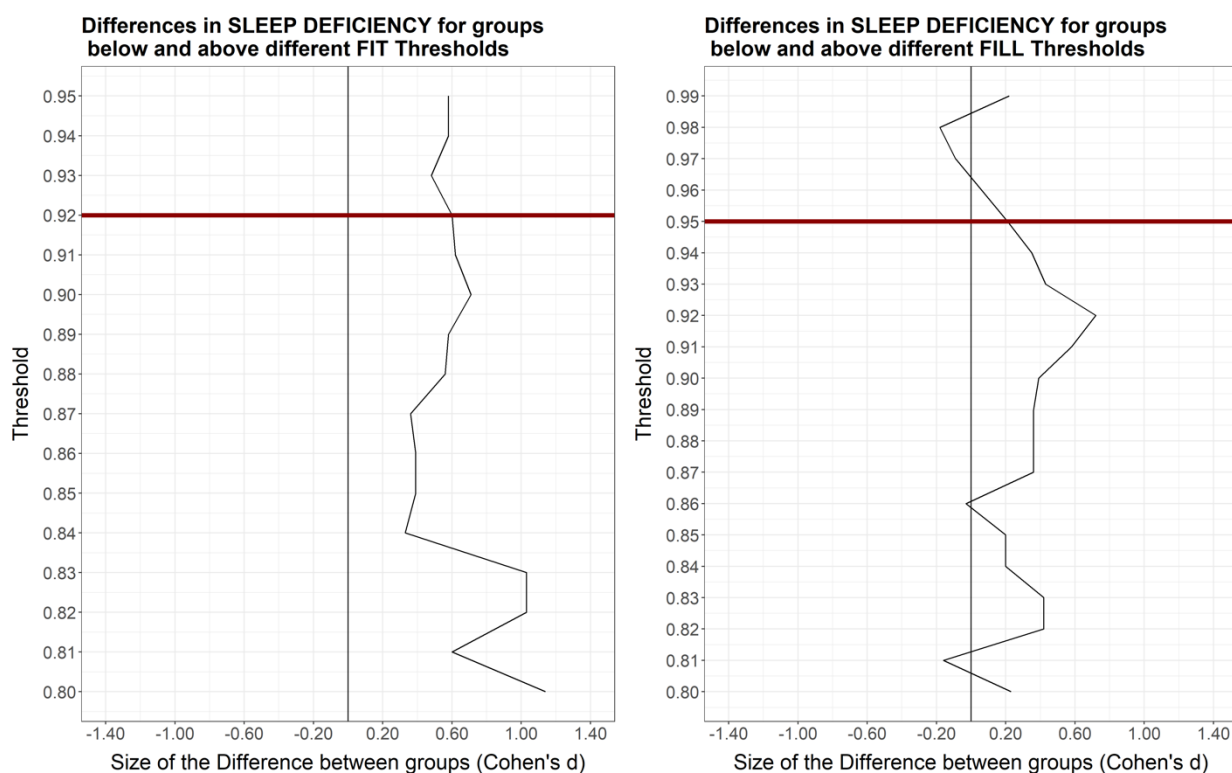


Figure 34 Comparison of Sleep Deficiency across different Fit and Fill Criterion Thresholds



4.7.2.5 Workload as a Factor Affecting Sleep

“Workload” refers to personnel beliefs about the degree to which workload affects their sleep, with ratings ranging from 1 (*not at all*) to 5 (*extremely*). Manning levels, if they are inadequate, may force greater work burdens on personnel, which could plausibly compromise sleep. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group implicating workload less often than the bottom group as a cause of sleep issues.

Figure 25 displays Cohen’s *d*, with negative values indicating that the subgroup satisfying the threshold exhibits less of the undesired response (i.e., implicating workload as a cause of sleep issues). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to workload (Cohen’s *d* $\approx$ 0). Further, values of Cohen’s *d* $<$ 0 also indicate the most plausible direction.

For the FIT criterion, all differences are in the expected direction (all Cohen’s *d* values $<$ 0), but none of the differences are statistically significant ($p > 0.05$). Values near the current FIT threshold of 0.92 were suggestive of nontrivial magnitudes (0.90 $\rightarrow$ -0.45; 0.91 $\rightarrow$ -0.41; 0.92 $\rightarrow$ -0.29), but there is insufficient evidence to suggest that these are stable differences. For the FILL criterion, thresholds at or below 0.94 are in the expected direction, but none of the differences displayed are statistically significant. Curiously, the trend reverses at the FILL threshold of 0.95 (the current threshold), but these differences are insufficiently reliable to warrant further scrutiny.

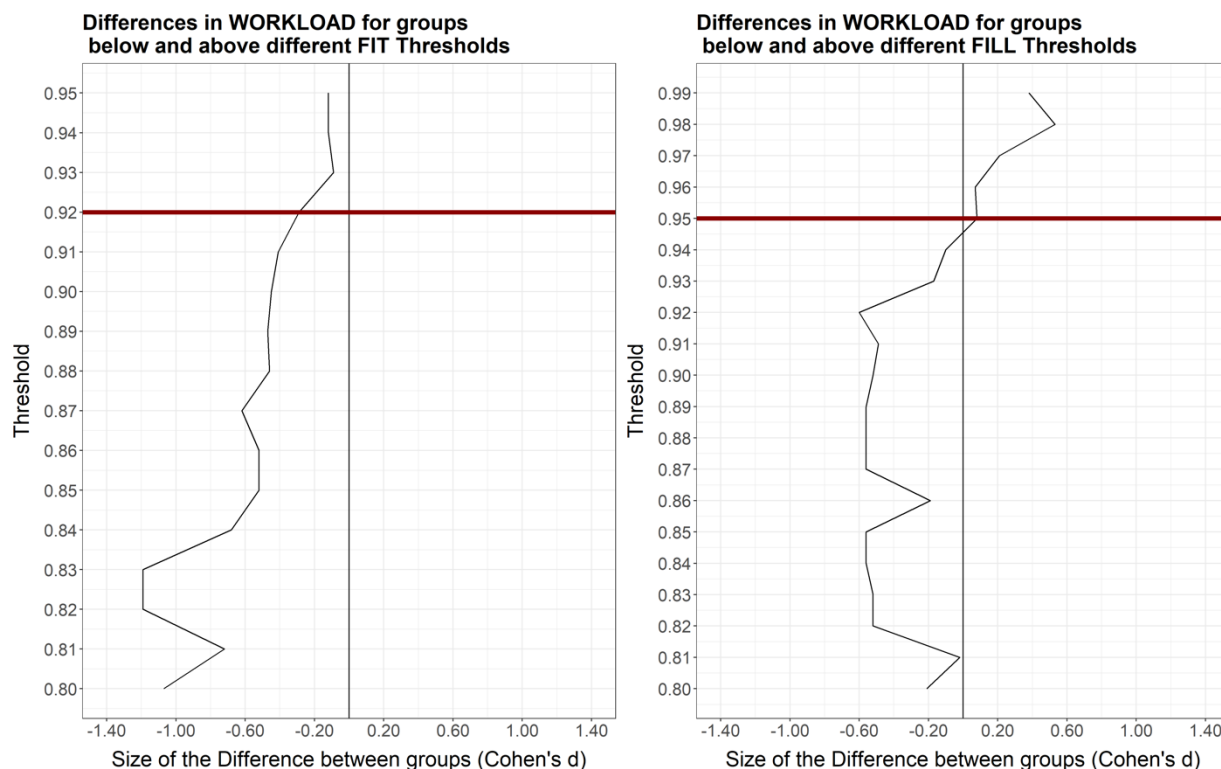


Figure 35 Comparison of Workload Across Different Fit and Fill Criterion Thresholds



4.7.2.6 Impairment Due to Fatigue

Impairment due to fatigue is quantified by a composite score -- a sum across 4 items -- with larger values indicating more frequent impairment due to fatigue. Manning levels, if they are inadequate, may force greater work burdens on personnel, which could plausibly increase fatigue, either in terms of disrupting restorative sleep or increasing time-on-task fatigue. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group exhibiting less fatigue than the bottom group.

Figure 26 displays Cohen's d , with negative values indicating that the subgroup satisfying the threshold exhibits less of the undesirable outcome (i.e., less fatigue). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to the amount of experienced fatigue (Cohen's $d \approx 0$). Further, values of Cohen's $d < 0$ indicate the most plausible direction.

For the FIT criterion, all differences are in the expected direction (all Cohen's d values < 0), with thresholds from 0.92 to 0.95 reaching statistical significance ($p < 0.05$). Of these values, the thresholds of 0.93 and 0.92 produced the smallest differences (-0.74 and -0.72, respectively), with the other thresholds producing medium-sized difference (≈ 0.60). For the FILL criterion, all of the thresholds (0.81, 0.86, 0.97, 0.98) produced differences in the expected direction, but none of the differences displayed are statistically significant. However, for FILL at 0.95 (the current threshold), there is a suggestive, nontrivial difference ($d = -0.39$); and for FILL at 0.92, there is a potentially stronger relationship ($d = -0.75$, $p = 0.10$). For now, however, there is insufficient evidence to suggest that either difference is stable ($p > .05$).

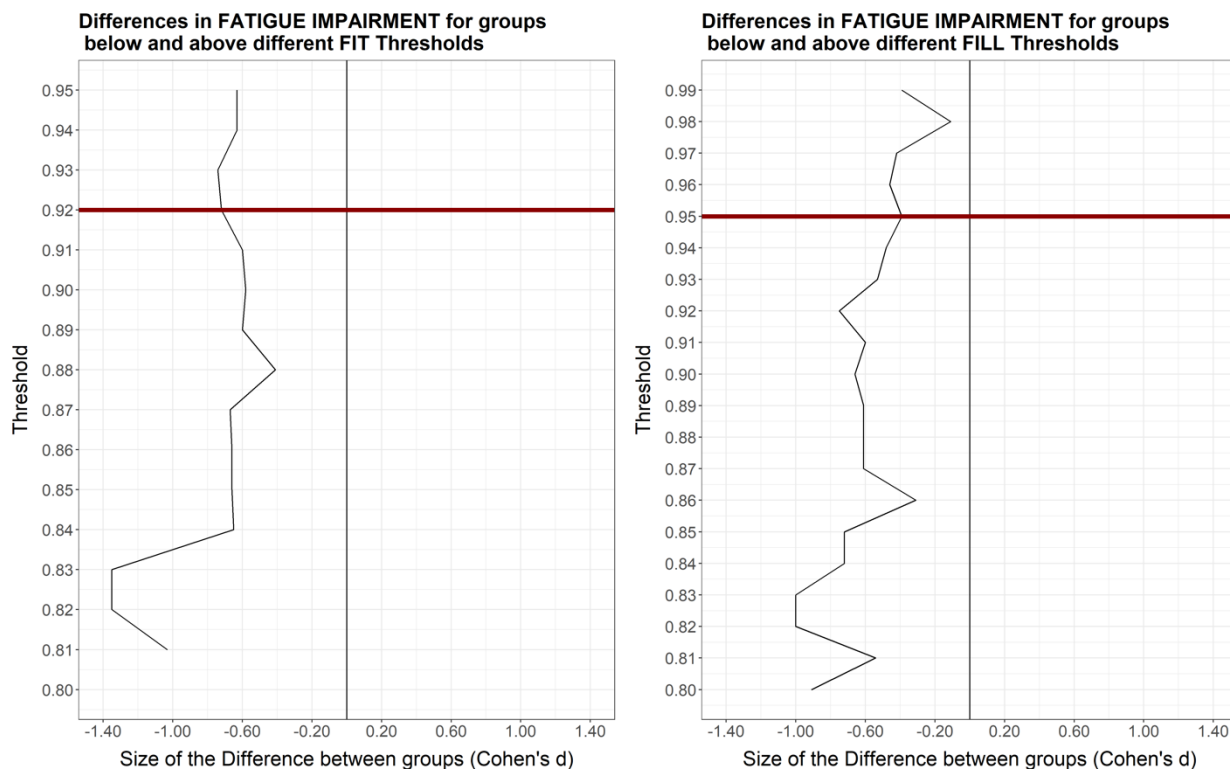


Figure 36 Comparison of Fatigue Impairment Across Different Fit and Fill Criterion Thresholds



4.7.2.7 Sleep Disturbance

“Sleep disturbance” refers, informally, to the extent to which one’s sleep is challenging and erratic (e.g., challenging to fall asleep, remain asleep, or for sleep to be restorative). It is quantified by a composite score—a sum across 4 items—with larger values indicating greater disturbance. Manning levels, if they are inadequate, may force greater work burdens on personnel, which could plausibly compromise the restorative aspects of sleep. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group exhibiting less disturbance than the bottom group.

Figure 27 displays Cohen’s d , with negative values indicating that the subgroup satisfying the threshold exhibits less of the undesirable outcome (i.e., less disturbance). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to the amount of disturbance (Cohen’s $d \approx 0$). Further, values of Cohen’s $d < 0$ indicate the most plausible direction.

For the FIT criterion, all but two thresholds (0.94 and 0.95) produce differences in the expected direction (all Cohen’s d values < 0), though none of these differences are statistically significant ($p > 0.05$). The current criterion threshold of 0.92 produces a small-to-medium size difference (-0.38), which is the largest absolute magnitude compared to its neighboring threshold values. For the FILL criterion, all but one of the thresholds (0.81) produce differences in the expected direction, but none of the differences displayed are statistically significant. However, for FILL at 0.95 (the current threshold), there is a suggestive, nontrivial difference ($d = -0.23$); and for FILL at 0.92, there is a potentially stronger relationship ($d = -0.45$, $p = 0.23$). For now, however, there is insufficient evidence to suggest that either difference is stable ($p > .05$).

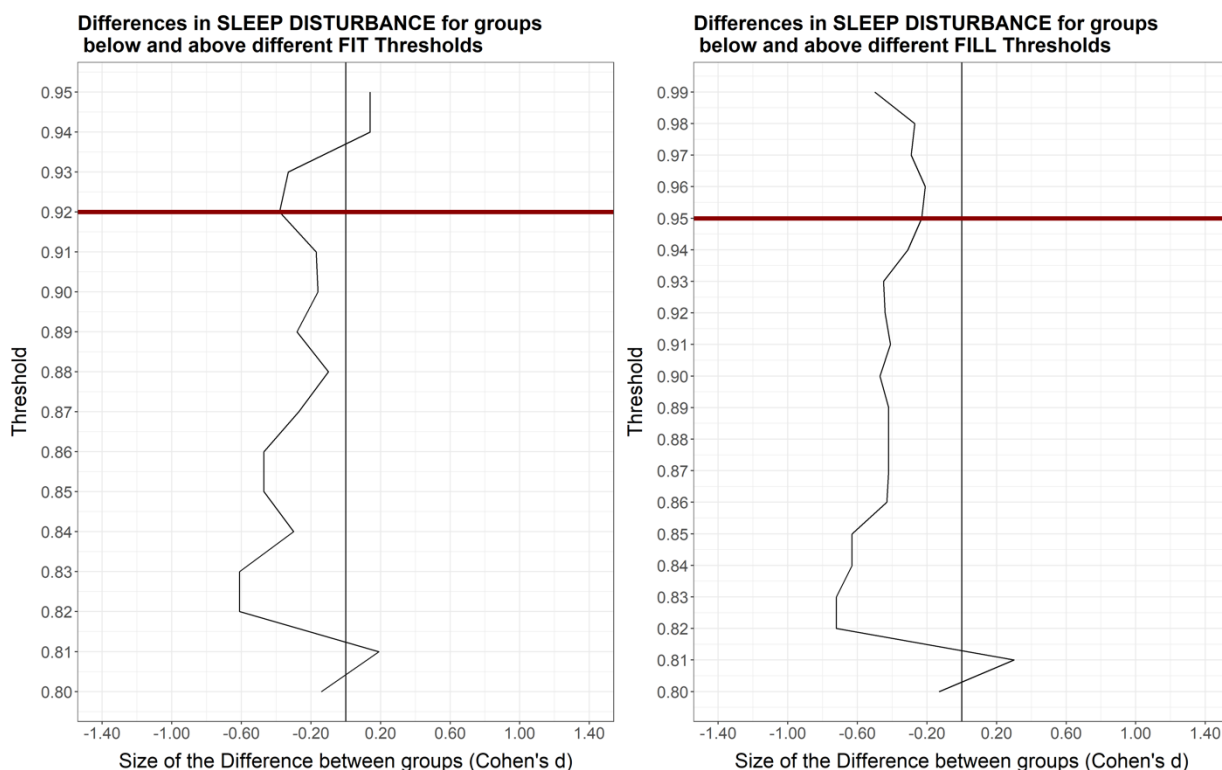


Figure 37 Comparison of Sleep Disturbance Across Different Fit and Fill Criterion Thresholds



4.7.2.8 Refreshing Sleep

As noted above, “sleep disturbance” refers, informally, to the extent to which one’s sleep is challenging and erratic (e.g., challenging to fall asleep, remain asleep, or to be restorative). One component -- the extent to which sleep is judged to be restorative -- is of interest on its own. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group exhibiting more restorative, refreshing, sleep than the bottom group (or equivalently, the bottom group experiences less restorative sleep).

Figure 28 displays Cohen’s d , with negative values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., more restorative, refreshing sleep). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to the amount of refreshing sleep (Cohen’s $d \approx 0$). Further, values of Cohen’s $d < 0$ indicate the most plausible direction (the scale has been reversed so that larger values reflect ratings of less restorative sleep).

For the FIT criterion, all but two thresholds (0.89 and 0.90) produce differences in the expected direction (all Cohen’s d values < 0), though none of these thresholds are statistically significant ($p > 0.05$). The current criterion threshold of 0.92 produces a small-to-medium size difference (-0.33), which is the largest absolute magnitude compared to its neighboring threshold values, though, again, it is not statistically significant. For the FILL criterion, all thresholds produce differences in the expected direction, but none of the differences displayed are statistically significant. However, for FILL at 0.95 (the current threshold), there is a suggestive, nontrivial difference ($d = -0.19$); and for FILL at 0.92, there is a potentially much smaller difference ($d = -0.53$, $p = 0.22$). For now, however, there is insufficient evidence to suggest that either difference is stable ($p > .05$).

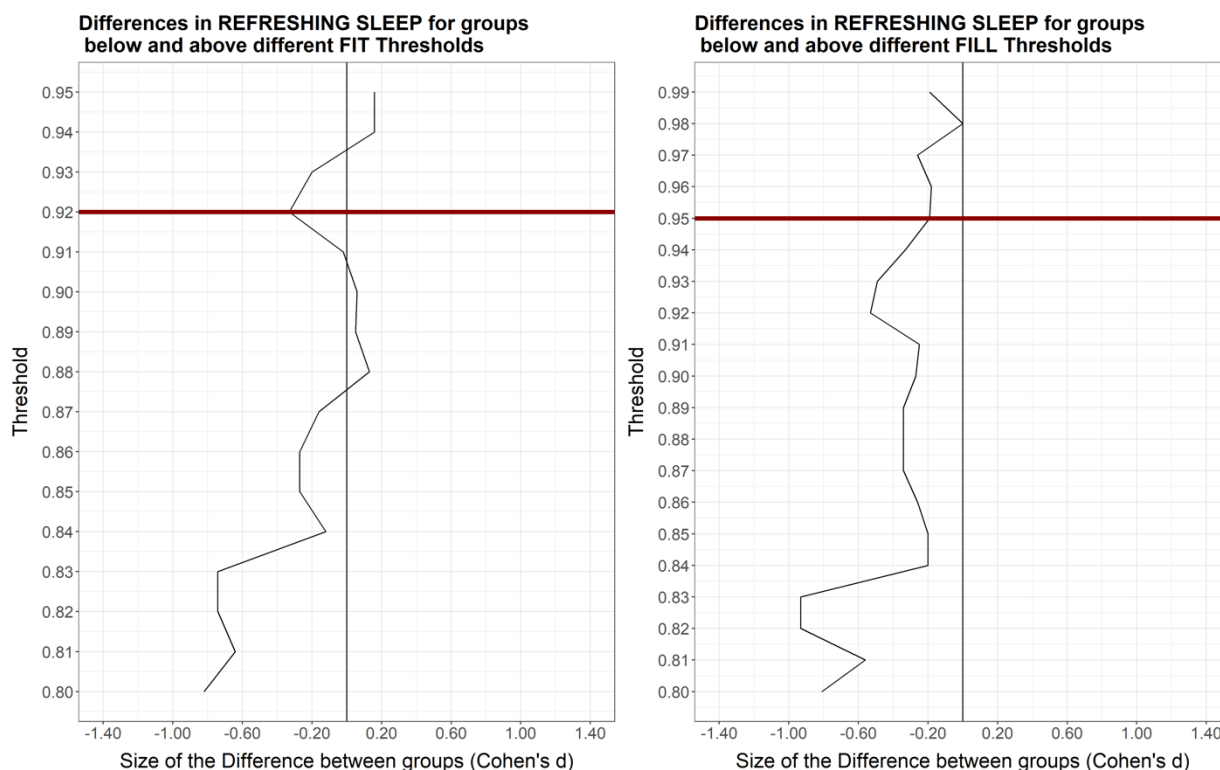


Figure 38 Comparison of Refreshing Sleep Across Different Fit and Fill Criterion Threshold.



4.7.2.9 Time Available to Exercise

Insufficient manning could plausibly lead to greater work burden, and less personal time to engage in important activities for maintaining well-being, such as exercise time. The time available to exercise is based on judgments of whether adequate time is provided for exercise. A useful criterion and defensible choice of threshold will separate the sample into two very different groups, with the top group exhibiting more time for exercise than the bottom group.

Figure 29 displays Cohen's d , with positive values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., more time for exercise). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to the amount of exercise time (Cohen's $d \approx 0$). Further, values of Cohen's $d > 0$ indicate the most plausible direction.

For the FIT criterion, all but three thresholds (0.84, 0.85, and 0.86) produce differences in the expected direction (all Cohen's d values > 0), though only the thresholds of 0.93, 0.94, and 0.95 are statistically significant ($p < 0.05$). The current criterion threshold of 0.92 produces a trivial difference (0.13), but its nearest neighbor (.93) produces the largest separation of 0.64, which is a medium-to-large difference. For the FILL criterion, only the thresholds at or above 0.95 produce differences in the expected direction, but none of the differences displayed are statistically significant. The FILL threshold of 0.95 produces a trivial difference ($d = 0.08$), and that of 0.98 produces the largest difference ($d = 0.53$), but there is insufficient evidence to suggest that the latter difference is stable ($p > .05$). It is important to note, however, that no mean ever fell below 3 for any group, which is the marker for "neutral" opinion.

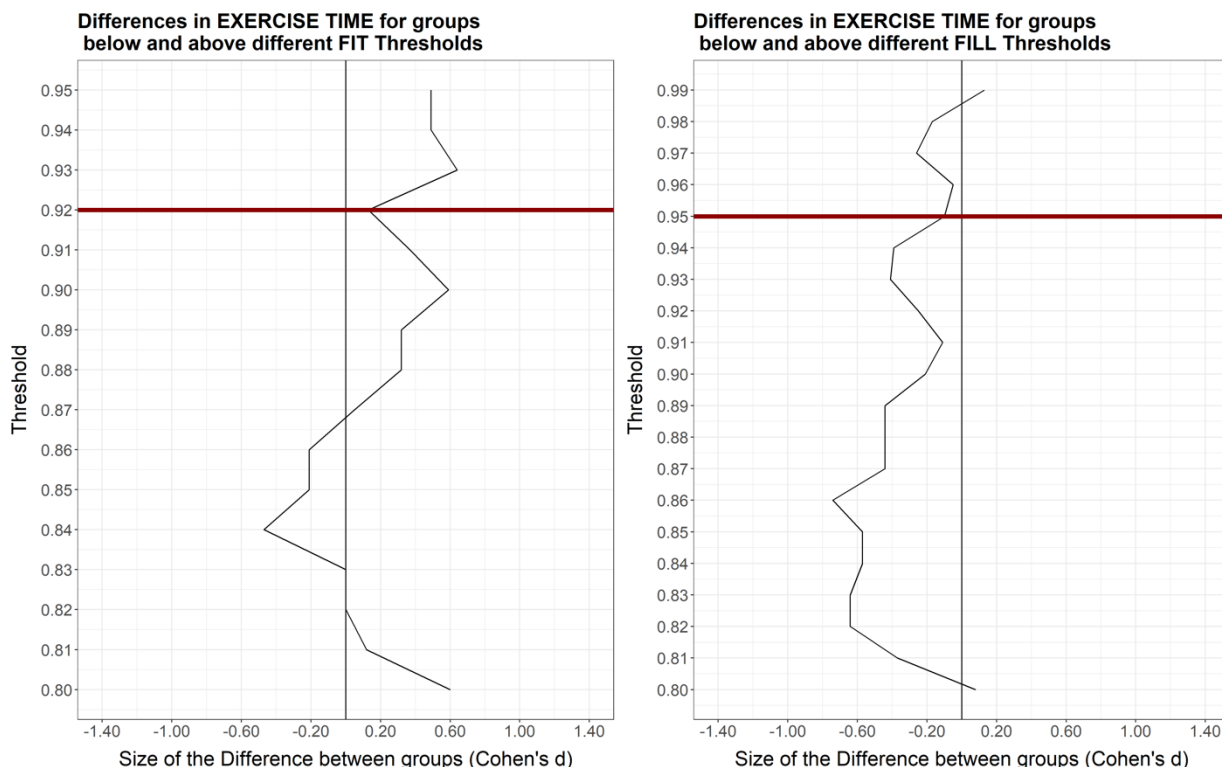


Figure 39 Comparison of Exercise Time Across Different Fit and Fill Criterion Thresholds



4.7.2.10 Leadership Assessment

Leadership assessment is captured by a composite score -- a sum across 4 items -- with larger values indicating more favorable views of leadership capabilities. Manning levels, if they are inadequate, may force greater work burdens on personnel, which could plausibly increase frustration, with leadership identified as one source of such frustration. Figure 30 displays Cohen's d , with positive values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., more positive views of leadership). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect leadership assessment (Cohen's $d \approx 0$). Further, values of Cohen's $d > 0$ indicate the most plausible direction.

For the FIT criterion, all but three thresholds (0.84, 0.85, 0.86, and 0.88) produce differences in the expected direction (all Cohen's d values > 0), though only the threshold of 0.92 is statistically significant ($p < 0.05$). The current criterion threshold of 0.92 produces the largest difference in the two groups ($d = 0.89$). For the FILL criterion, only the thresholds at or above 0.95 produce differences in the expected direction, but none of the differences displayed are statistically significant. Of note, the FILL threshold of 0.95 produces a suggestive, nontrivial difference ($d = 0.22$), but again, it is not statistically significant.

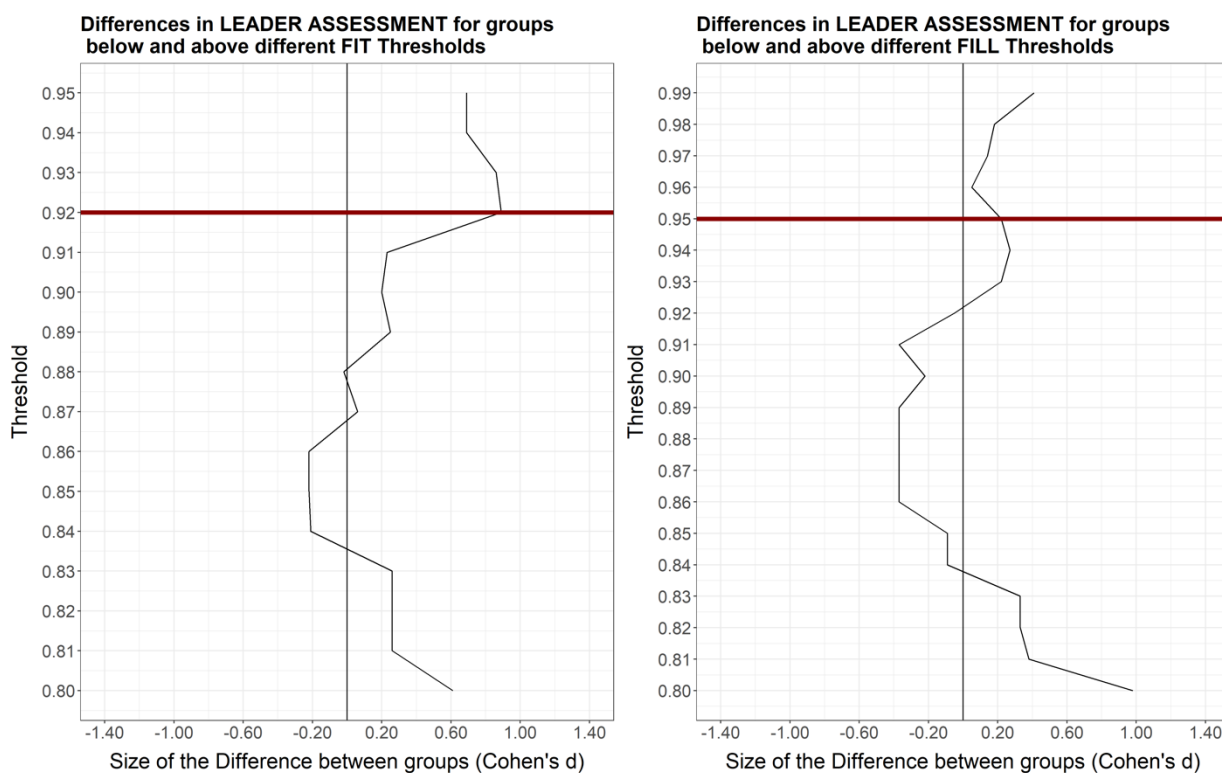


Figure 40 Comparison of Leadership Assessment Across Different Fit and Fill Criterion Thresholds



4.7.2.11 Affective Commitment

Affective commitment attempts to capture the extent to which an individual is emotionally connected to his or her command. It is measured by a composite score -- a sum across 3 items -- such that larger values indicate stronger emotional connection. Manning levels, if they are inadequate, may force greater work burdens on personnel, which could plausibly increase frustration, which could lead to a breakdown in any emotional ties to the command. Figure 31 displays Cohen's d , with positive values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., stronger emotional connect). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect leadership assessment (Cohen's $d \approx 0$). Further, values of Cohen's $d > 0$ indicate the most plausible direction.

For the FIT criterion, all thresholds at or greater than 0.89 produce differences in the expected direction (all Cohen's d values > 0), though none are statistically significant ($p > 0.05$). The current criterion threshold of 0.92 does produce, however, a suggestive nontrivial difference ($d = 0.59$, $p = 0.12$), which is second in size only to the FIT criterion threshold of 0.94 ($d = 0.63$). For the FILL criterion, most differences (all but thresholds of 0.96 and 0.99) are in the unexpected opposite direction: that is, the bottom group tends to express *stronger* affective commitment, though the differences are statistically significant only for 0.84 and 0.85 ($d = -1.03$, $p = 0.039$ for both). This result points to an interesting possibility that should continue to be explored as more data comes in.

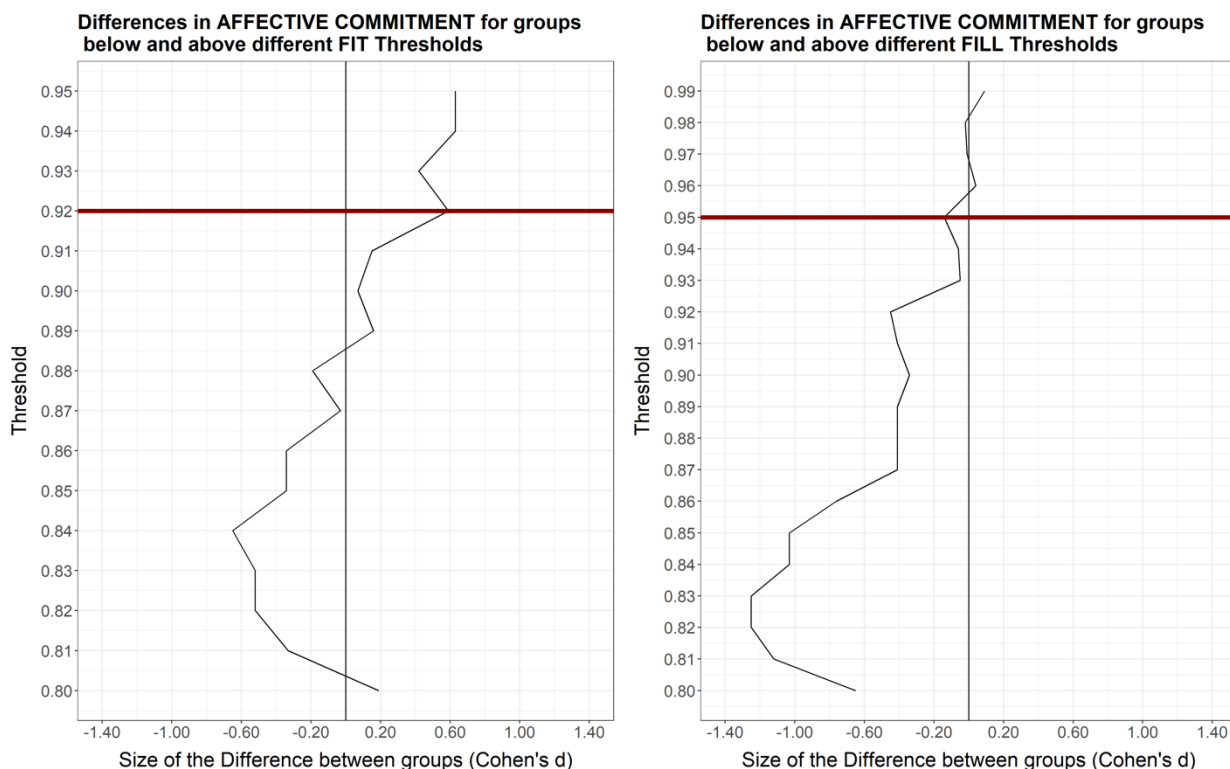


Figure 41 Comparison of Affective Commitment Across Different Fit and Fill Criterion Thresholds



4.7.2.12 Command Satisfaction

Command satisfaction is measured by a composite score -- a sum across 2 items -- that reflects the extent to which personnel are happy with, and motivated to perform best, for the command. Manning levels, if they are inadequate, may force greater work burdens on personnel, which could increase frustration, which could plausibly undermine command satisfaction. Figure 32 displays Cohen's d , with positive values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., greater command satisfaction). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect leadership assessment (Cohen's $d \approx 0$). Further, values of Cohen's $d > 0$ indicate the most plausible direction.

For the FIT criterion, all but three thresholds (0.81, 0.82, 0.83, and 0.84) produce differences in the expected direction (all Cohen's d values > 0). However, only the threshold of 0.92 produces a statistically significant, and the largest, difference ($d = 0.69$, $p < 0.05$). For the FILL criterion, only the thresholds at or above 0.90 produce differences in the expected direction, but none of the differences displayed are statistically significant. Of note, the FILL threshold of 0.95 produces a suggestive, nontrivial difference ($d = 0.29$), while the threshold of 0.93 produces the largest difference ($d = 0.38$), but again, neither is statistically significant.

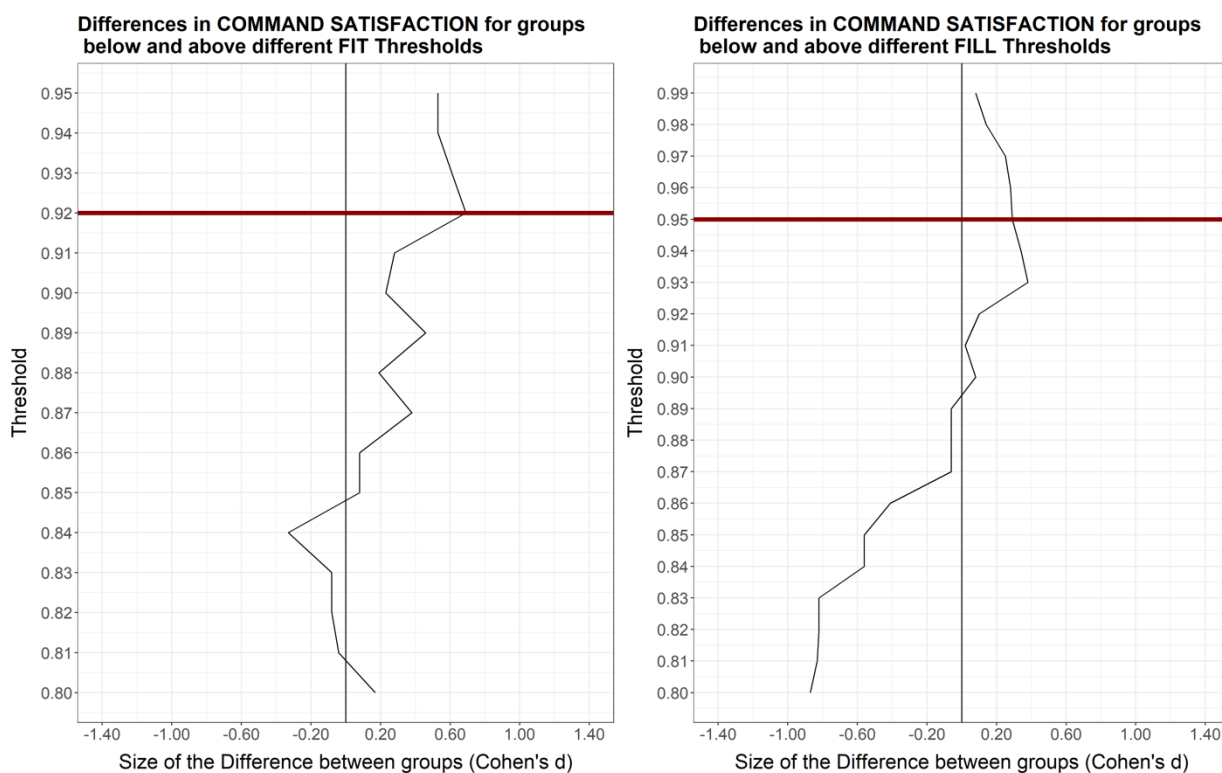


Figure 42 Comparison of Command Satisfaction Across Different Fit and Fill Criterion Thresholds



4.7.2.13 Command Safety Practices

Command safety practices are quantified as single composite score -- a sum across 4 items -- that reflects the extent to which personnel judge the command as promoting a culture of safety. Manning levels, if they are inadequate, may force greater work burdens on personnel, which could increase the presence of threats to safety. Figure 33 displays Cohen's d , with positive values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., greater expression of command safety practices). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect leadership assessment (Cohen's $d \approx 0$). Further, values of Cohen's $d > 0$ indicate, arguably, the most plausible direction.

For the FIT criterion, all thresholds at or greater than 0.89 produce differences in the expected direction (all Cohen's d values > 0), though none are statistically significant ($p > 0.05$). The current criterion threshold of 0.92 does produce, however, a suggestive nontrivial difference ($d = 0.37$, $p = 0.42$), which is second in size only to the FIT criterion threshold of 0.93 ($d = 0.44$, $p = .40$). For FILL thresholds, all but one (0.99) produce differences in the *opposite* direction: the bottom group expressing more positive judgments of the command's efforts to promote a culture of safety. Indeed, only those thresholds at or below 0.83 produces differences that are statistically significant, and quite large in size ($ds < -1.70$). These provisional results may suggest that leadership recognizes the shortage in occupied positions and engages in efforts to compensate. This is a hypothesis that should be explored as more data continue to come in. Note: for the FILL figure below, values that extend past the lower boundary on the x-axis are not displayed; they are provided, however, in the table.

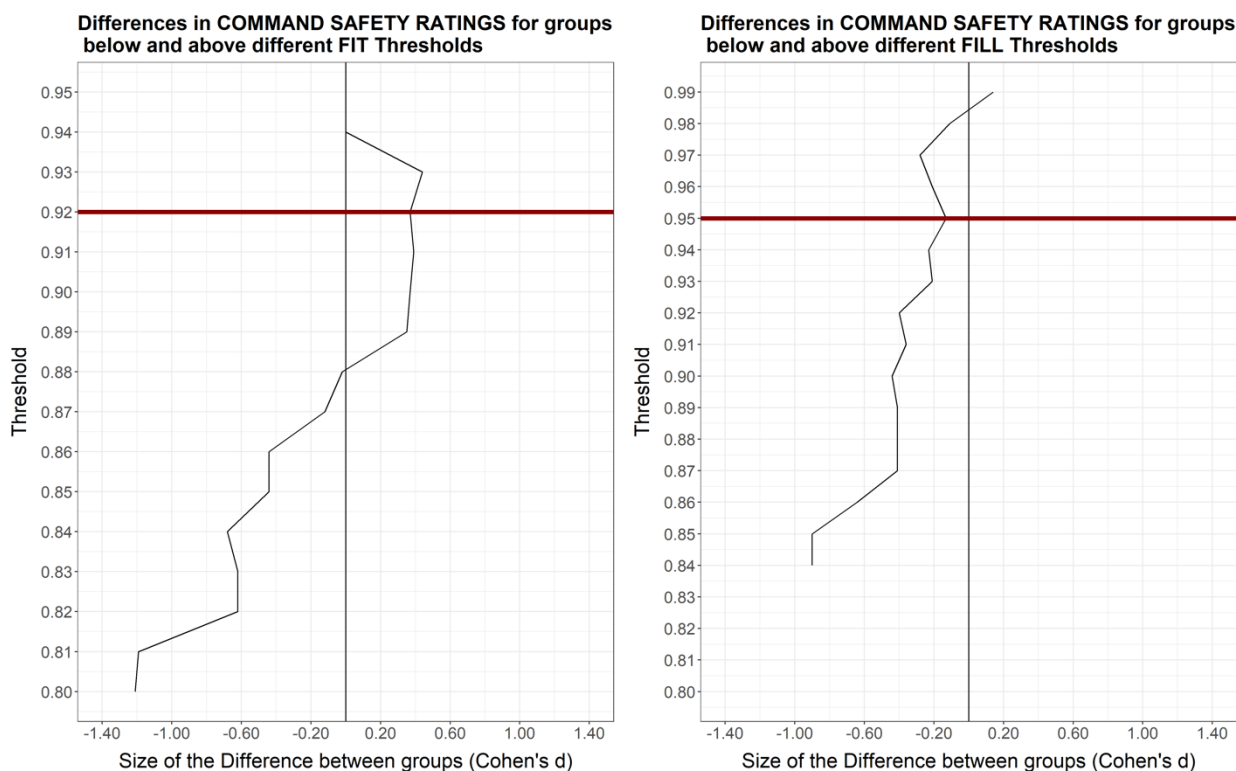


Figure 43 Comparison of Command Safety Ratings Across Different Fit and Fill Criterion Thresholds



4.7.2.14 Social Support

Social support is quantified and represented as a single composite score -- a sum across 4 items -- and is intended to capture the strength and extent of a given individual's access to social resources, possibly as a mitigation strategy against the effects of fatigue and heavy workload. Figure 34 displays Cohen's d , with positive values indicating that the subgroup satisfying the threshold exhibits more of the desired property (i.e., greater access to social support). Values close to the vertical line indicate that the top and bottom groups are indistinguishable with respect to social support (Cohen's $d \approx 0$). Further, values of Cohen's $d > 0$ indicate, arguably, the most plausible direction. An alternative hypothesis, however, is that those in the bottom group face greater challenges, and thus the need for social support is greater. The understaffed command may understand this and will commit resources towards cultivating supportive social networks.

For the FIT criterion, there is little evidence that any thresholds produce noteworthy separation in groups. For FILL, there are interesting trends that could potentially suggest that those in the bottom group have greater access to social support resources, especially for those environments that are particularly understaffed.

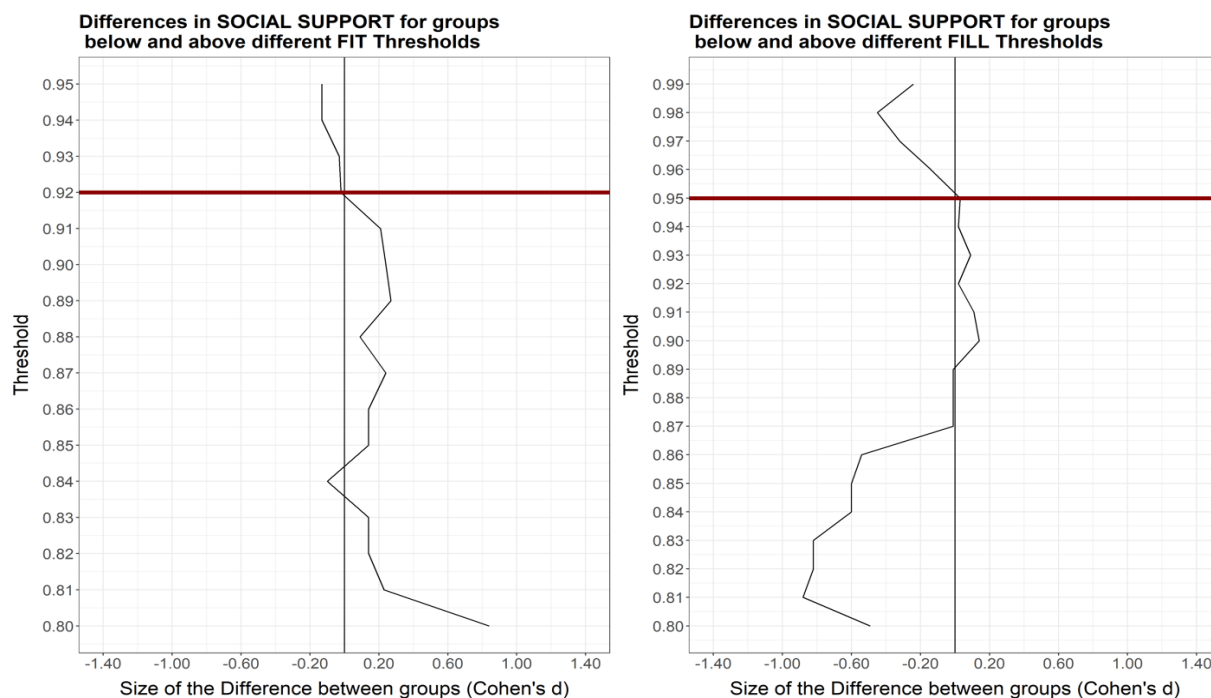


Figure 44 Comparison of Social Support Across Different Fit and Fill Criterion Thresholds



4.8 OFRP Analysis

The ASCAS instructions require platforms to complete their surveys during the READ-E1 (Basic) and READ-E5 (Sustainment) phases of the OFRP. Analysis was performed to identify whether or not phase correlated with aggregate responses of the constructs of interest amongst 13 crews in the Basic-phase and 20 in the Sustainment-phase. Table 21 shows the differences between these means and their tests of significance and effect size.

Table 18: Mean Differences Between 1(Basic) and 5 (Sustainment)

OFRP Phase	OFRP Phase				
	Phase 1	Phase 5	Mean Difference	P-Value	Effect Size
Sleep Disturbance *	3.31	3.34	0.03	0.21	-0.02
Sleep Factors *^	2.78	2.88	0.10	<0.01	-0.09
Sleep Impairment *	2.40	2.43	0.03	0.03	-0.02
Job Stress *	3.55	3.67	0.12	<0.01	-0.11
Job Demands *	3.08	3.29	0.21	<0.01	-0.18
General Health *^	2.96	2.81	-0.14	<0.01	0.15
Mental Health *+	10.01	11.23	1.22	<0.01	-0.11
Physical Health *+	4.96	6.01	1.06	<0.01	-0.12

Note*: For these questions lower scores represent a more favorable response

Note *+ : For these questions, the values represent the Number of Days. They are also not constructs; therefore, a Cronbach Alpha calculation would not be relevant.

Note *^ : This question is not a construct; therefore, a Cronbach Alpha calculation would not be relevant.

Regarding OFRP phase, those in RE1 (returning from deployment) had significantly better scores on Sleep Factors, Job Stress, Mental and Physical than RE5 (transitioning from the maintenance to the basic phase), but significantly poorer scores on General Health.

4.9 Natural Language Processing (NLP) Analysis on Open-Ended Comments

4.9.1 SENTIMENT ANALYSIS

To better understand the open-ended comments, we performed a sentiment analysis of the free response field. Sentiment Analysis is an NLP technique that can determine whether text is positive, negative, or neutral in connotation. Since the text has many Navy-specific terms and phrases that might not be accurately assessed by any single sentiment analysis method, an ensemble of three different methods of varying complexity was developed and used. The overall sentiment score reported here is an average of the three methods: TextBlob, vaderSentiment, and Flair. The score is scaled to be between 0 and 1, with 0 corresponding to negative sentiment. Of the 3,313 Sailors who opted-in to the survey, 1,079 provided open-ended responses. The distribution of the sentiment scores is presented by paygrade, department, and Ship Class.

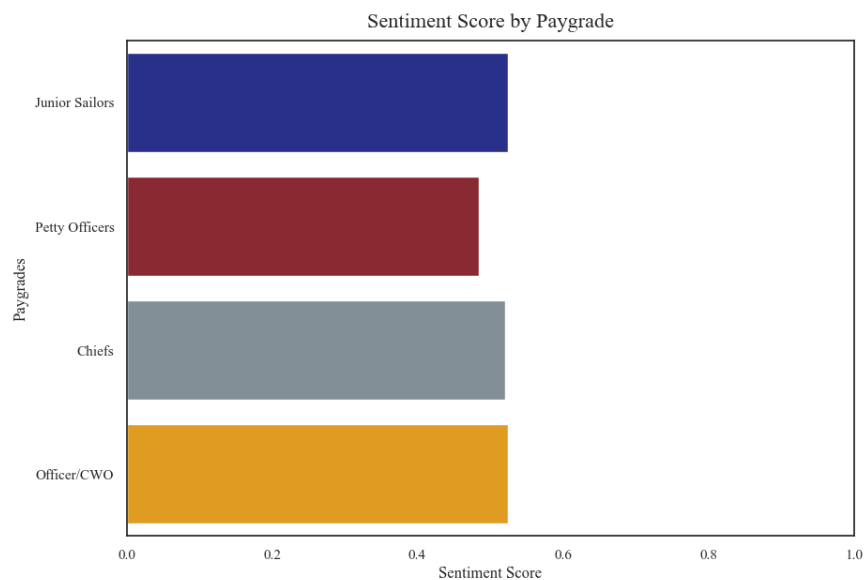


Figure 45: Bar Chart of Sentiment Score by Paygrade

The main trend to note in Figure 35 is that there is a steady decrease in the number of negative comments going from Petty Officers to Chiefs to Officers/CWOs, although there is not necessarily an increase in positive comments along those lines.

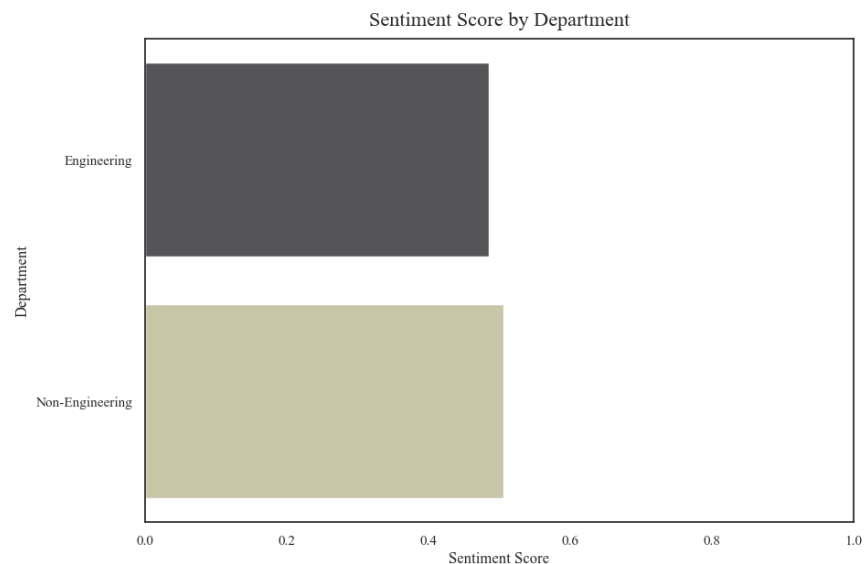


Figure 46: Bar Chart of Sentiment Score by Department

Based on the quantitative survey data, it was hypothesized that those in the Engineering department would have a greater percentage of negative comments than those in other departments. When we break down the sentiment score by department in Figure 36, we see that the medians are nearly identical. There is, however, a definite skew toward negative comments from those serving in Engineering.

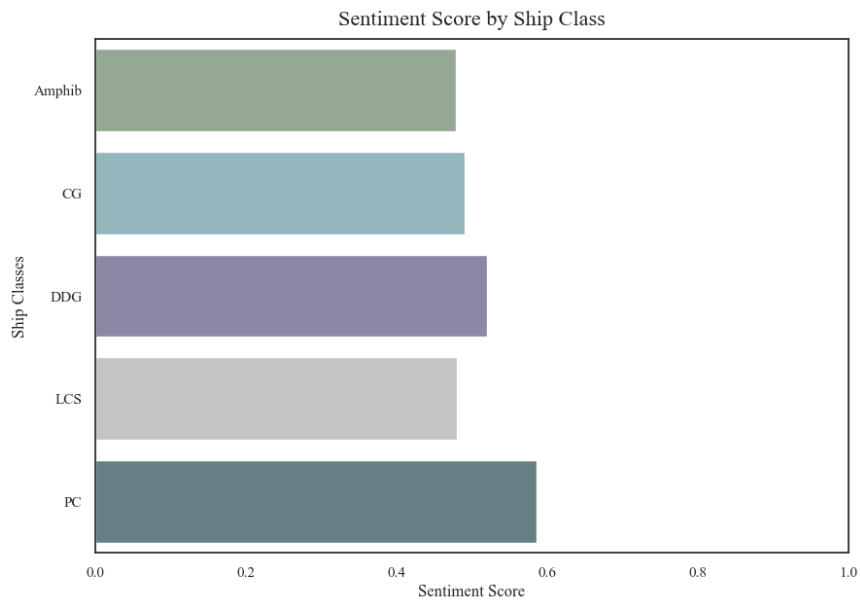


Figure 47: Bar Chart Sentiment Score by Ship Class

The main takeaway from Figure 37 is that there is much more negative sentiment from Sailors serving on LCSs. DDGs have a sentiment that skews slightly positive compared to the others. The smaller range of the distribution of sentiment scores for the PC class is due to there being only 14 open-ended comments.

4.9.2 CLUSTER ANALYSIS

Analysis was performed to identify topics within the open-ended comments. While these results are preliminary, they show clusters related to crew endurance, manning, training, nutrition, lack of sleep, lack of exercise, workload, morale, and mental health. To achieve these results, open-ended comments were processed, vectorized, clustered and interpreted. Note that there are three open-ended response questions (overall feedback, best safety practices and worst safety practices) in the survey. The table below provides topic summaries and sample responses from each cluster.



Table 19: Results of Cluster Analysis

Cluster Analysis			
Cluster ID	Topics	Sample Response	Number of Responses
0	Morale, Manning, Sleep	"The crew has been running on fumes since the year began. What is the light at the end of the tunnel? There will always be more work down the road. When will the crew have the time to focus on themselves again?"	98
1	Sleep	"There needs to be a better way for sailors to report that they aren't getting enough sleep so it can be investigated. Often my shipmates or myself have tried to report that we aren't getting enough sleep and it is always overlooked. Deck department has constantly taken away sleep hour for watch standers and free time from workers."	67
2	Covid-19, Leadership, Sleep, Mental Health, Manning	"Manning will continue to play a huge role in how we manage safety and health concerns for the crew. COVID has put a great deal of strain on our Sailors and with the continued loss of PLs or extension of orders, they will continue to struggle with hitting even the bare minimum and will be likely to take more risks for the sake of getting things done. The sooner the Navy goes back to (what will hopefully be an improved) 'normal' the better!"	41
3	Manning, Sleep, Watchstanding, Safety, Leadership, Equality, Workload	"The problems I see are not exclusive to this ship, but common across ships. Too many Navy instituted programs, inspections, or tasks are individually important and accomplishable, but in combination with each other require cutting corners to get all of them done. Corners such as safety, sleep, or procedure. They Navy never seems to accept failure as an option."	466
4	Mental health, low morale	"I don't know who is missing the big picture here, but the navy overall has a big mental health issue with its sailors. There needs to be more thought and consideration on the mental health of the crew than there is at this moment in time."	52
5	Scheduling concerns	"Departmental meetings should be held at a time convenient with one's schedule so that time one's supposed to be asleep isn't interrupted."	22
6	Sleep, nutrition	"Food onboard is disgusting and feels like there is no care being put into the food. Often times things are overcooked and not seasoned at all. The ship does circadian rhythm but with so few crew members we are in 3 section and the sleep rotation for 3 section is really bad. On duty days we muster at 0530 in the morning to clean and the sailors who have the 2200-0200 watch are required to be there so they do not get adequate sleep after standing watch. And most times there is no chance to sleep before watch because the whole duty section gets mustered for light detail."	44
7	None	"NTR"	17
8	Nutrition	"The Ship store needs more nutritional snacks, too much junk food not good for overall health. Not enough portions being given during chow, More portions of food should be given out during chow hours especially dinner."	20
9	Crew endurance, self-care, manning, training	"Perhaps the most significant issues can be solved with increased manning. The lack of trained personnel causes greatly extended work hours underway and in port."	72
10	Crew endurance, exercise, Nutrition	"Offer counseling sessions to the crew regarding sleep, exercise and nutrition."	32
11	None	"No comment"	13
12	None	"N/A"	19



13	Exercise	“It would be great if we could be allowed time to exercise. As it stands we have no group exercise schedule and relieve very little free time to exercise.”	3
14	Manning, Sleep	“Every division onboard is undermanned putting greater stress on divisions and leadership when it comes to assessments and inspections as well as just trying to maintain the daily work list with undermanned work centers. By being undermanned it incentives working in unsafe conditions to meet deadlines.”	38
15	Exercise, Nutrition	“Carve out time in the workday for PT. Yes, we have 90 minutes for lunch, so it is possible to sneak in a workout then, but it would be great if there were designated time in the morning or late afternoon for this, even once per week.”	7

Note: many clusters included responses on a variety of topics. Larger clusters tend to have a wider array of topics. Clusters 7, 11, and 12 captured open-ended comments without topics and were removed from subsequent analyses. While this unsupervised ML technique allows for automated grouping and quantification of resulting classes, results are preliminary and should be reviewed and refined by SMEs.

Analysis was performed to identify themes and distributions of clusters by various demographics. To allow for ease of interpretation, clusters were aggregated into the following groupings:

Table 20: K-mean Clusters Aggregated into Topics

Topics	Cluster IDs	Count
Workload/Manning/Safety/Sleep/Equality/Misc.	3	466
Sleep/Nutrition/Exercise	1, 6, 8, 10, 11, 13, 14, 15	211
Morale/Manning/Sleep	0	98
Crew Endurance/Manning/Training	9	72
Morale/Mental Health	4	52
Covid-19/Misc.	2	41
Scheduling	5	22

Cluster analysis was performed in comparison to paygrade as reported in Table 25.

Table 21: Cluster Analysis by Paygrade

Paygrade	Number of Open-ended Comments Responses	Number of Survey Respondents	Percentage of Respondents (%)
E1-E3	108	474	22.78
E4-E6	578	1838	31.44
E7-E9	111	363	30.58
Officers/CWOs	164	524	31.30

Representative samples for each paygrade are included in the analysis. Notably, fewer E1-E3 respondents provided open-ended responses (22.78%) compared to the E4-E6 (31.44%), E7-E9 (30.58%) and Officer/CWOs (31.30%). Figure below details where the various Paygrade comments fall within the clusters.

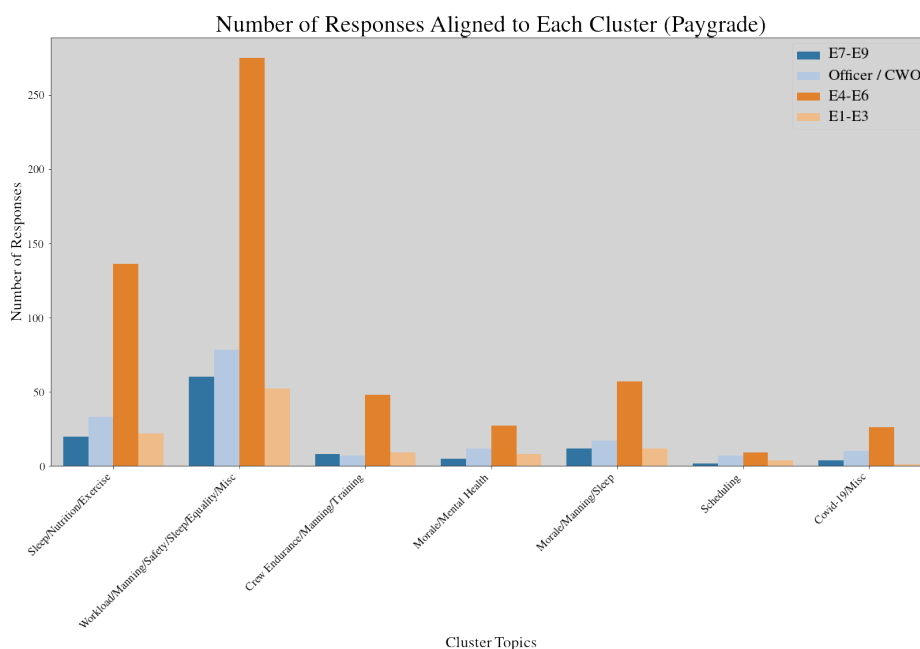


Figure 48: Number of Responses Aligned to Each Cluster by Paygrade

In all groups, the most responses were in the “Workload/Manning/Safety/Sleep/Equality/Misc.” cluster. While E7-E9s had noticeably fewer responses related to scheduling and E1-E3 respondents tended to have fewer responses aligned to the “Covid-19/Misc.” cluster.

Table 22: Cluster Analysis by Ship Class

Ship Class	Number of Open-ended Responses	Number of Survey Respondents	Percentage of Respondents (%)
DDG	425	1395	30.47
Amphibs	316	1168	27.05
CG	137	452	30.31
LCS	73	255	28.63
PC	11	43	25.59

With regard to Ship Class, the percentage of respondents who left open-ended comments was highest for DDGs (30.47%) and CGs (30.31%).

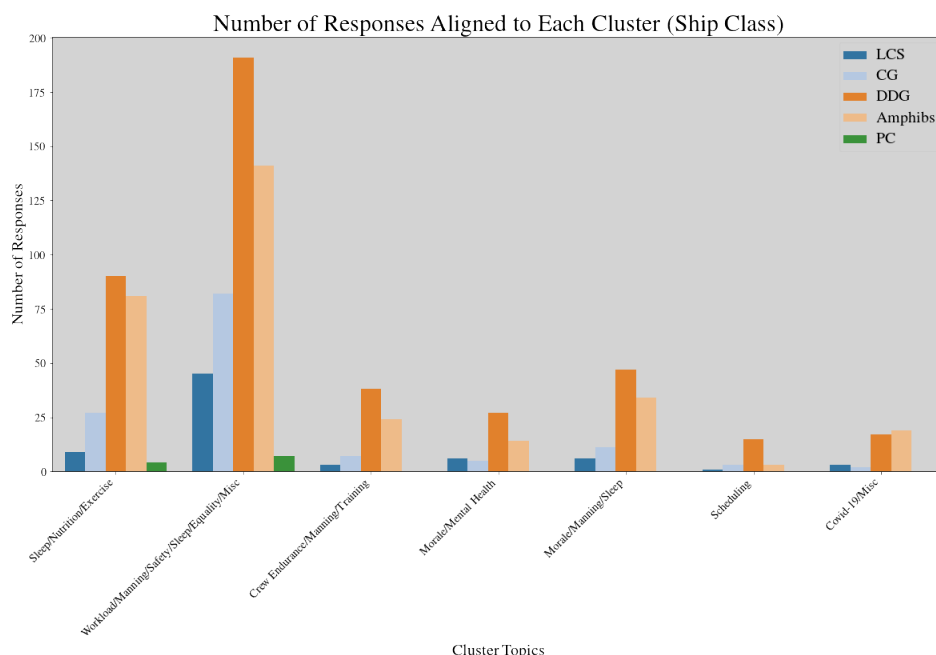


Figure 49: Number of Responses Aligned to Each Cluster by Ship Class

Similar to paygrade, the “Workload/Manning/ Safety/Sleep/Equality/Misc.” cluster had the most responses across all Ship Classes. DDGs, CGs, Amphibs, and LCS’s responses spread across the different clusters, with the most other responses aligned to “Sleep/Nutrition/Exercise” and “Morale/Manning/Sleep” concerns. The PC comments were mainly limited to “Workload/Manning/Safety /Sleep/Equality/Misc.” and “Sleep/Nutrition/Exercise”. Notably, Amphibs had more comments aligned to the “Covid-19/Misc.” than all other groups. LCSs had more “Morale/ Mental Health” comments when compared to other distributions. More granular insight can be gleaned when visualizing responses at the individual ship level.



5 CONCLUSION

5.1 Summary

The ASCAS is an insightful resource to capture a comprehensive operational safety profile within SURFOR. More broadly, the ASCAS can be viewed as part of a larger effort that seeks to: 1) identify crew problems; and 2) provide robust data that can later be incorporated into analyses that attempt to identify the underlying causes of those problems (e.g., using structural equation modeling or Bayesian network construction, both of which can support “What if?” tests of the possible outcomes from implementing different courses of action). Both components are essential for successfully monitoring operational safety.

The conclusions drawn in this report are provisional and may change as more ASCAS data are collected; 35 additional ships have already completed the survey. Of particular importance, with a larger sample size, the OSRI team will be in a better position to generate norms for judging when a problem is present or absent in a given command. As it stands now, in defining an “area of praise or concern,” preliminary thresholds of .05 above and below the “neutral” option served to identify problematic or praiseworthy circumstances. However, this choice of criterion is limited as it does not take into account the individual variation that might hold across constructs. With that caveat in mind, overall, SURFOR’s areas of praise included Leader Perception of Subordinates, Unit Cohesion (Work Center and Watch Team), Supervisor Safety Climate, Personal Responsibility for Workplace Safety, Command Safety Practices, Sleep Factors and Job Resources (namely Equipment). Areas of concern included Affective Commitment, Satisfaction of Command, Job Stress, Job Demands, Team Processes, Psychological Safety and No-Blame Error Reporting. Mean differences between demographic groups in the sample were typically negligible, though women and older respondents reported fewer days of poor mental health.

When comparing CNSP and CNSL, one fleet did not consistently respond more positively than the other, and any observed differences were typically negligible. Regardless, CNSP expressed slightly more favorable responses towards Affective Commitment, Error Management, Command Noncompliance, and Satisfaction with Command; while CNSL expressed slightly more favorable responses towards Job Stress, Occupational Challenges, Job Demands, Days with not good mental health, and Days with not good physical health. There is, however, insufficient evidence of any noteworthy differences in safety climate between CNSP and CNSL.

Of greater note are the differences observed in paygrade. These analyses identified differences in nearly every survey construct, with Enlisted Sailors responding more unfavorably than Officers/CWOs. Constructs manifesting the largest differences included Unit Cohesion (Watch Team), Leader Assessment, Affective Commitment, Error Management, Psychological Safety, Safety Satisfaction, Personal Responsibility for Workplace Safety, Satisfaction with Command, Days which mental health was judged ‘not good’, and Days which physical health was judged ‘not good’.

Similar to paygrade outcomes, the analyses highlighted notable differences in Sailors aligned to the Engineering department compared to other ship departments. In particular, Sailors assigned to Engineering departments reported more unfavorable responses across the constructs of Job Stress, Occupational Challenges, Safety Satisfaction, Job Demands, Sleep Impairment, Days of poor mental health and Days of poor physical health.

The ASCAS year one dataset was joined to several other metrics to better understand how these factors affect safety climate. Specifically, FIT/FILL metrics and OFRP phase data were assessed. Commands with below standard FILL (< 95%) did not reveal any statistically significant differences compared to those where FILL > 95%. However, platforms with low FIT (<92%) reflected notably less



favorable responses in the following areas: Leader Assessment, Affective Commitment Construct, Team Process, Crew Resilience Construct, Error Management, No-Blame Error Reporting, Psychological Safety, Safety Satisfaction, Job Demands, Days of poor mental health and days of poor physical health. OFRP analyses indicated that Sailors in the Sustainment-phase had slightly more days of poor mental and physical health than those in the Basic-phase.

To cultivate an understanding of those opinions solicited from the open-ended responses, NLP techniques were applied to quantify sentiments and to classify comment topics. The data revealed a suggestive trend of more negative sentiment displayed by Sailors on LCSs than in other Ship Classes. The open-ended comments covered a range of topics, with clusters related to the survey topics of crew endurance, manning levels, training issues, nutrition issues, lack of sleep, lack of exercise, high workload, low morale, and mental health concerns.

These findings serve to help SURFOR identify and mitigate operational safety and crew endurance concerns. The ASCAS lays the foundation that will ultimately help promote a deeper and richer understanding of the complex multifaceted interplay of factors that contribute to overall safety culture with a high-risk organization. The Navy is poised to leverage robust analytic techniques for use on the ASCAS data, which will produce a tool that will complement ongoing strategic lines of effort, ultimately enabling better decision making and more effective implementation of countermeasures designed to promote safety.

5.2 Next Steps

During FY22, the OSRI will apply more sophisticated analytical techniques as the ASCAS dataset grows. Additionally, the ASCAS data is being used with other disparate datasets such as personnel systems (e.g., FLT MPS, COGNOS, NT MPS, and DMDC), training (e.g., STRMS and TORIS), health-related records (e.g., NMCPHC and ADMITS), safety-related data (RMI and SURFOR Critiques and Lessons Learned), and other data sources (e.g., 90-day SURFOR CO reports) in order to commence ML-based modeling. Latent factor model techniques such as principal components analysis, exploratory factor analysis, and structural equation modeling will serve to identify plausible hypotheses about underlying factors impacting operational safety, human performance, and general readiness. Additionally, different regression methods such as multilevel modeling are being used to help account for the clustering effects of ships, thus enabling better control over responses that may be highly correlated within commands.



6 APPENDICES:

6.1 Appendix A: Ship Class Additional Analyses

Comparison across Ship Classes was comprehensive. Section 4.4 details the most striking findings. The Tables below provide insight to additional distinguishing measures between the groups.

Table 23: Ship Type Areas of Praise and Concern by Ship Class

Item	Destroyer Mean (Likert: 1-5)	Cruiser Mean (Likert: 1-5)	Amphib Mean (Likert: 1-5)	Littoral Combat Ship Mean (Likert: 1-5)
Leader Perception of Subordinates (Subset 3)	3.97	3.79	3.82	4.09
Leadership Perception of Subordinates (Subset 2)	3.79	3.64	3.62	3.88
Unit Cohesion (Watch Team)	3.75	3.68	3.58	3.97
Unit Cohesion (Work Center)	3.70	3.63	3.45	3.89
Leadership Perception of Subordinates (Subset 1)	3.63	3.58	3.54	3.73
Supervisor Safety	3.60	3.54	3.53	3.75
Personal Responsibility for Workplace Safety	3.57	3.49	3.54	3.88
Command Safety Practices	3.50	3.42	3.48	3.64
Job Resources (Equipment)	3.49	3.22	3.42	3.49
Safety Satisfaction	3.37	3.34	3.31	3.61
Job Stress *	3.72	3.57	3.42	3.79
Affective Commitment	2.72	2.66	2.39	3.01
No-Blame Error Reporting	2.79	2.93	2.76	3.08
Job Demands *	3.28	3.18	3.06	3.31
Team Process (Interpersonal Processes)	2.92	2.90	2.73	3.08
Psychological Safety	2.94	2.95	2.79	3.21
Team Process (Action Processes)	2.98	2.91	2.86	3.17
Safety Reporting Behavior *	3.00	3.03	2.98	2.78
Team Processes (Overall)	3.00	2.95	2.84	3.15
Satisfaction of Command	3.02	2.96	2.80	3.11
Team Process (Transition Processes)	3.10	3.07	2.94	3.17
Sleep Disturbances *	3.31	3.38	3.36	3.47

Table 24: Tests of Significance and Effect Size for Sleep Disturbance by Ship Class

Sleep Disturbance			
Ship Type	Mean Difference	P-Value	Effect Size
CG-Amphib	0.05	0.63	-0.04
DDG- Amphib	-0.03	0.84	0.02
LCS-Amphib	0.14	0.01	-0.12
PC-Amphib	-0.13	0.69	0.11
DDG-CG	-0.07	0.18	0.06
LCS-CG	0.09	0.3	-0.08
PC-CG	-0.17	0.41	0.14
LCS-DDG	0.16	0	-0.14
PC-DDG	-0.1	0.83	0.08
PC-LCS	-0.27	0.08	0.22



Table 25: Tests of Significance and Effect Size for Sleep Impairment by Ship Class

Sleep Impairment			
Ship Type	Mean Difference	P-Value	Effect Size
CG-Amphib	0.05	0.41	-0.05
DDG- Amphib	0.10	<0.01	-0.09
LCS-Amphib	0.15	<0.01	-0.14
PC-Amphib	0.01	1.00	-0.01
DDG-CG	0.04	0.61	-0.04
LCS-CG	0.10	0.17	-0.09
PC-CG	-0.04	0.99	0.04
LCS-DDG	0.05	0.60	-0.05
PC-DDG	-0.09	0.87	0.08
PC-LCS	-0.14	0.56	0.13

Table 26: Tests of Significance and Effect Size for Sleep Factors by Ship Class

Sleep Factors			
Ship Type	Mean Difference	P-Value	Effect Size
CG-Amphib	0.03	0.73	-0.02
DDG- Amphib	-0.10	<0.01	0.07
LCS-Amphib	-0.25	<0.01	0.18
PC-Amphib	-0.30	<0.01	0.22
DDG-CG	-0.13	<0.01	0.09
LCS-CG	-0.27	<0.01	0.20
PC-CG	-0.33	<0.01	0.24
LCS-DDG	-0.14	<0.01	0.10
PC-DDG	-0.20	<0.01	0.15
PC-LCS	-0.06	0.88	0.04

Table 27: Tests of Significance and Effect Size for Job Stress by Ship Class

Job Stress			
Ship Type	Mean Difference	P-Value	Effect Size
CG-Amphib	0.05	0.59	-0.05
DDG- Amphib	0.20	<0.01	-0.19
LCS-Amphib	0.27	<0.01	-0.25
PC-Amphib	0.22	0.21	-0.20
DDG-CG	0.15	<0.01	-0.14
LCS-CG	0.22	<0.01	-0.20
PC-CG	0.16	0.52	-0.15
LCS-DDG	0.07	0.55	-0.06
PC-DDG	0.01	1.00	-0.01
PC-LCS	-0.05	0.99	0.05



Table 28: Tests of Significance and Effect Size for Job Demands by Ship Class

Job Demands			
Ship Type	Mean Difference	P-Value	Effect Size
CG-Amphib	0.01	1.00	-0.01
DDG- Amphib	0.11	<0.01	-0.09
LCS-Amphib	0.14	<0.01	-0.12
PC-Amphib	-0.03	1.00	0.03
DDG-CG	0.10	0.01	-0.08
LCS-CG	0.13	0.02	-0.11
PC-CG	-0.04	0.99	0.03
LCS-DDG	0.03	0.93	-0.03
PC-DDG	-0.14	0.51	0.12
PC-LCS	-0.17	0.37	0.14

Table 29: Tests of Significance and Effect Size for Mental Health by Ship Class

Mental Health			
Ship Type	Mean Difference	P-Value	Effect Size
CG-Amphib	-0.26	0.98	0.03
DDG- Amphib	-0.10	1.00	0.01
LCS-Amphib	-0.59	0.86	0.07
PC-Amphib	1.18	0.91	-0.14
DDG-CG	0.15	1.00	-0.02
LCS-CG	-0.33	0.99	0.04
PC-CG	1.43	0.85	-0.17
LCS-DDG	-0.48	0.92	0.06
PC-DDG	1.28	0.88	-0.15
PC-LCS	1.76	0.75	-0.21

Table 30: Tests of Significance and Effect Size for General Health by Ship Class

General Health			
Ship Type	Mean Difference	P-Value	Effect Size
CG-Amphib	-0.49	0.93	0.05
DDG-Amphib	0.22	0.99	-0.02
LCS-Amphib	-0.5	0.96	0.05
PC-Amphib	-1.33	0.94	0.12
DDG-CG	0.71	0.75	-0.07
LCS-CG	-0.01	1	0
PC-CG	-0.85	0.99	0.08
LCS-DDG	-0.72	0.87	0.07
PC-DDG	-1.55	0.9	0.14
PC-LCS	-0.84	0.99	0.08

*Table 31: Tests of Significance and Effect Size for Physical Health by Ship Class*

Physical Health			
Ship Type	Mean Difference	P-Value	Effect Size
CG-Amphib	-0.05	0.85	0.06
DDG-Amphib	-0.08	0.25	0.08
LCS-Amphib	-0.08	0.8	0.08
PC-Amphib	0.08	0.99	-0.08
DDG-CG	-0.02	0.99	0.02
LCS-CG	-0.02	1	0.02
PC-CG	0.13	0.93	-0.13
LCS-DDG	0	1	0
PC-DDG	0.15	0.86	-0.16
PC-LCS	0.15	0.89	-0.16



6.2 Appendix B: Paygrade Mean Differences

Section 4.5 details the findings amongst paygrades. It is worth noting that means differed significantly between Officers/CWOs and Enlisted Sailors. Table 36 details all s and t-test results with differences greater than or equal to 0.15. Tables 37-44 capture analysis performed by paygrade.

Table 32: Means with Largest Differences based on Officer and Enlisted and Tests of Significance

Item	Officer			Enlisted			T-test		
	Mean	SD	Alpha	Mean	SD	Alpha	Diff	P-Value	Effect Size
Unit Cohesion (Work Center)	3.94	0.78	0.94	3.56	0.99	0.92	0.38	<0.001***	
Unit Cohesion (Watch Team)	4.04	0.77	0.93	3.62	0.94	0.94	0.42	<0.001***	
Leader Assessment	3.64	0.83	0.76	3.18	0.88	0.76	0.46	<0.001***	
Occupational Challenges	2.61	0.95	0.65	2.78	1.00	0.74	-0.17	0.002*	
Affective Commitment	3.12	1.09	0.91	2.53	1.09	0.91	0.59	<0.001***	
Social Support	3.42	0.98	0.90	3.16	1.04	0.91	0.26	<0.001***	
Team Process	3.21	0.78	0.95	2.88	0.81	0.95	0.33	<0.001***	
Crew Resilience	3.32	0.86	0.87	3.02	0.93	0.90	0.3	<0.001***	
Error Management	3.66	0.82	0.69	3.26	0.89	0.70	0.4	<0.001***	
No-Blame Error Reporting	3.06	0.72	0.68	2.77	0.76	0.70	0.29	<0.001***	
Psychological Safety	3.30	0.73	0.76	2.83	0.81	0.79	0.57	<0.001***	
Supervisor Safety Climate	3.83	0.78	0.91	3.53	0.90	0.90	0.3	<0.001***	
Command Safety Practices	3.73	0.72	0.88	3.42	0.80	0.91	0.31	<0.001***	
Individual Noncompliance	2.52	0.91	0.82	2.78	0.91	0.85	-0.26	<0.001***	
Safety Satisfaction	3.69	0.76	0.88	3.27	0.89	0.90	0.42	<0.001***	
Personal Responsibility for Workplace Safety	3.98	0.76	0.87	3.49	0.86	0.86	0.49	<0.001***	
Safety Communication	3.61	0.81	0.90	3.24	0.87	0.89	0.37	<0.001***	
Safety Reporting Behavior	2.83	0.84	0.80	3.02	0.85	0.85	-0.19	<0.001***	
Job Resources	3.58	0.66	0.88	3.35	0.74	0.90	0.33	<0.001***	
Satisfaction with Command	3.53	1.07	0.81	2.81	1.17	0.84	0.72	<0.001***	
Sleep Impairment	2.64	0.83	0.88	2.89	0.95	0.89	-0.25	<0.001***	
Days which mental health was judged 'not good' *	7.57	8.89		11.42	10.92		-3.85	<0.001***	
Days which physical health was judged 'not good'*	4.02	7.07		5.98	8.79		-1.96	<0.001***	



Table 33: Tests of Significance and Effect Size for Sleep Disturbance by Paygrade

Sleep Disturbance			
Paygrade	Mean Difference	P-Value	Effect Size
Petty Officers-Junior Sailors	0.24	0.00	-0.20
Chiefs-Junior Sailors	0.14	0.00	-0.12
Officers/CWOs-Junior Sailors	-0.25	0.00	0.21
Chiefs-Petty Officers	-0.10	0.02	0.08
Officers/CWOs-Petty Officers	-0.49	0.00	0.41
Officers/CWOs-Chiefs	-0.39	0.00	0.33

Table 34: Tests of Significance and Effect Size for Sleep Factors by Paygrade

Sleep Factors			
Paygrade	Mean Difference	P-Value	Effect Size
Petty Officers-Junior Sailors	0.06	0.01	-0.04
Chiefs-Junior Sailors	-0.30	0.00	0.22
Officers/CWOs-Junior Sailors	-0.42	0.00	0.31
Chiefs-Petty Officers	-0.36	0.00	0.26
Officers/CWOs-Petty Officers	-0.49	0.00	0.36
Officers/CWOs-Chiefs	-0.13	0.00	0.09

Table 35: Tests of Significance and Effect Size for Sleep Impairment by Paygrade

Sleep Impairment			
Paygrade	Mean Difference	P-Value	Effect Size
Petty Officers-Junior Sailors	0.18	0.00	-0.16
Chiefs-Junior Sailors	-0.08	0.16	0.07
Officers/CWOs-Junior Sailors	-0.14	0.00	0.13
Chiefs-Petty Officers	-0.26	0.00	0.24
Officers/CWOs-Petty Officers	-0.32	0.00	0.29
Officers/CWOs-Chiefs	-0.06	0.34	0.06

Table 36: Tests of Significance and Effect Size for Job Stress by Paygrade

Job Stress			
Paygrade	Mean Difference	P-Value	Effect Size
Petty Officers-Junior Sailors	0.25	0.00	-0.23
Chiefs-Junior Sailors	0.25	0.00	-0.23
Officers/CWOs-Junior Sailors	0.28	0.00	-0.26
Chiefs-Petty Officers	0.00	1.00	0.00
Officers/CWOs-Petty Officers	0.03	0.72	-0.03
Officers/CWOs-Chiefs	0.03	0.91	-0.03



Table 37: Tests of Significance and Effect Size for Job Demands by Paygrade

Job Demands			
Paygrade	Mean Difference	P-Value	Effect Size
Petty Officers-Junior Sailors	0.16	0.00	-0.14
Chiefs-Junior Sailors	0.17	0.00	-0.14
Officers/CWOs-Junior Sailors	0.19	0.00	-0.16
Chiefs-Petty Officers	0.00	1.00	0.00
Officers/CWOs-Petty Officers	0.03	0.74	-0.02
Officers/CWOs-Chiefs	0.03	0.90	-0.02

Table 38: Tests of Significance and Effect Size for Mental Health by Paygrade

Mental Health			
Paygrade	Mean Difference	P-Value	Effect Size
Petty Officers-Junior Sailors	0.40	0.81	-0.05
Chiefs-Junior Sailors	0.72	0.62	-0.08
Officers/CWOs-Junior Sailors	-1.59	0.02	0.19
Chiefs-Petty Officers	0.33	0.91	-0.04
Officers/CWOs-Petty Officers	-1.99	0.00	0.23
Officers/CWOs-Chiefs	-2.31	0.00	0.27

Table 39: Tests of Significance and Effect Size for General Health by Paygrade

General Health			
Paygrade	Mean Difference	P-Value	Effect Size
Petty Officers-Junior Sailors	1.57	0.02	-0.15
Chiefs-Junior Sailors	-1.14	0.41	0.11
Officers/CWOs-Junior Sailors	-2.92	0.00	0.28
Chiefs-Petty Officers	-2.71	0.00	0.26
Officers/CWOs-Petty Officers	-4.49	0.00	0.42
Officers/CWOs-Chiefs	-1.78	0.07	0.17

Table 40: Tests of Significance and Effect Size for Physical Health by Paygrade

Physical Health			
Paygrade	Mean Difference	P-Value	Effect Size
Petty Officers-Junior Sailors	-0.08	0.33	0.09
Chiefs-Junior Sailors	0.12	0.28	-0.13
Officers/CWOs-Junior Sailors	0.46	0.00	-0.49
Chiefs-Petty Officers	0.20	0.00	-0.21
Officers/CWOs-Petty Officers	0.54	0.00	-0.57
Officers/CWOs-Chiefs	0.34	0.00	-0.36