What Do the Sailors Say? Using Natural Language Processing to Interpret Open-Ended Responses about Safety Issues

Adam T. Biggs,[1] Ryan Lance,[2] Todd R. Seech,[1] & Dale W. Russell[1,3]

[1] Commander, Naval Surface Force, US Pacific Fleet, Coronado, California, USA

[2] Booz Allen Hamilton

[3] Department of Psychiatry, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA

Running Head: SAILOR CONCERNS ABOUT SAFETY

Please address all correspondence to:

Dale W. Russell
Commander, Naval Surface Force, U.S. Pacific Fleet
2841 Rendova Road, Coronado, CA 92155
240-330-0038
dale.w.russell1.civ@us.navy.mil

Author Note

 This work was funded by Commander, Naval Surface Forces. Neither the authors nor their family members have a financial interest in any commercial product, service, or organization providing financial support for this research. The views expressed herein are those of the authors and do not necessarily reflect the official policy or position of the Department of Defense, Department of the Navy, Uniformed Services University nor the U.S. Government. The survey was approved by the Naval Medical Center Portsmouth Institutional Review Board (NMCP.2020.0051).

Abstract

Occupational safety is a critical component of organizational success, especially in industries with high-risk working conditions. Maximizing an organization's occupational safety outcomes requires routinely assessing its safety climate to help identify emerging issues as well as opportunities for improvement. Naval operations present particularly complex environments since working and living spaces are intertwined and personnel serve in prolonged high-risk operations. Although many safety climate assessments probe this environment with structured surveys, there is the opportunity for the survey measure itself to inadvertently bias answers through pre-determined responses. To address this issue, open-ended questions allow personnel to report safety-related concerns without inducing surveyor bias. The current study utilized natural language processing and topic modeling to explore open-ended safety answers from a large sample of naval personnel. Answers indicated that personnel and readiness were among the most frequently reported safety-related problems along with scheduled sleep and watchstanding. Moreover, senior leaders offered notably different comments about the nature of best and worst safety-related practices compared to junior personnel. This outcome demonstrates some common elements in the most prominent safety concerns in maritime operations while also noting some systemic differences in perceptions that are influenced by position within the leadership hierarchy.

Keywords: Safety; natural language processing; topic modeling; naval; endurance

1. Introduction.

Occupational safety impacts workplace operations in every industry, but especially for high-risk occupations, the consequences of failure can be exceptionally dangerous and costly (Bartone, & Barry, 2011). Effective safety procedures can be the difference between near misses, severe injuries, deaths—or, the ideal goal of no injuries and a positive working environment. Although each industry introduces unique complications, a subset of high-risk industries complicates the occupational environment further by having a shared living-working space. Maritime operations represent a prototypical example of this problem as sailors will live and work within the same ship hull for weeks or months at a time. In turn, there are few more complex examples of the substantial interdependent processes that can affect naval safety (Antão & Soares, 2019; Hänninen, 2014; Zhang & Thai, 2016). More importantly, recent events have indicated room for improvement given the frequency of devastating accidents in seafaring operations (Allianz, 2019; US Navy, 2017). These combined issues demonstrate the complexity and importance of identifying problems and enhancing naval safety.

Secure and reliable operations depend upon a collection of attitudes, perceptions, and best practices known as a safety climate (Neal & Griffin, 2002; 2004; Zohar, 1980; 2010; 2011). Specifically, this concept represents a myriad of safety-related attitudes that underscore effective working conditions. Even if the relationship between safety climate and safety outcomes remains complex (Alruqi, Hallowell, & Techera, 2018; Christian et al., 2009; Clarke, 2006; 2010), the most parsimonious conclusion is that safety climate is a multi-faceted concept related to injuries and health in the workplace. Still, among the many factors encompassed by safety climate, safety communication remains among the most important (Griffin & Neal, 2000). Safe working operations are dependent upon these communication behaviors. Personnel must feel free to

report problems for the leadership to have an accurate view of the safety climate within the organization. If communication breaks down, as with the pervasive problem of safety underreporting (Leigh et al., 2004; Lowery et al., 1998; Probst, 2015; Probst & Estrada, 2010; Rosenman et al., 2006), then there can be an escalating chance for unsafe practices to progress into smaller accidents and eventually major injuries (Bellamy, 2015; Bird & Germain, 1996; Heinrich, 1931).

Unfortunately, safety communication—and indeed many aspects of the safety climate—can be affected by other factors in the operating environment. A specific concern involves the relationship between organizational climate and safety climate. Whereas safety climate can be specific to safety-related practices, organizational climate encompasses a wider subset of occupational factors such as perceptions of leadership and affective commitment to the organization. Previous work has already demonstrated that organizational climate can be a powerful antecedent of safety climate (Neal, Griffin, & Hart, 2000; Zohar & Luria, 2005), yet this relationship can become even stronger in a naval operating environment given the shared living-working space. This conflation introduces other factors that could affect safe working conditions in a naval operating environment, including the likelihood of sleep issues aboard ship (Harrison et al., 2017; Russell, Markwald, & Jameson, 2021; Shattuck & Matsangas, 2016; 2017; 2022), berthing habitability problems (Jameson et al., 2022; Matsangas & Shattuck, 2017; 2021), noise exposure (Schaal, Lange, & Majar, 2019), and even the greater likelihood of physical injuries (Krentz, Li, & Baker, 1997; Schmidt, Schmorrow, & Figlock, 2000). In short, the organizational and safety climates aboard ship become nearly inseparable constructs that must be considered together when evaluating naval operations.

Given the multi-faceted nature of safety and organizational climates, there is the possibility that the act of measurement unintentionally biases the outcome. Researchers and practitioners may inadvertently omit certain variables or cause other variables to seem more important based upon the context in which they were observed or the overall set of items included when measuring organizational or safety climates. The possibility is akin to a Hawthorne effect (Roethlisberger & Dickson, 1967; Wickström & Bendix, 2000), where the act of observation itself influences the outcome. In this case, researchers and surveillance programs could influence the perceived importance of different factors based upon the variables they choose to include during the assessment process. An alternative is to utilize open-ended questions when evaluating safety climate issues. These constructs can be used to discover what an individual might provide spontaneously without being biased by suggesting possible responses (Reja et al., 2003). However, they also create problems related to scorer bias in coding answers and non-responses. These issues make open-ended questions a double-edged sword when trying to discover elements of the safety climate without introducing observer bias.

One effective solution to the open-ended problem involves a technique known as natural language processing (NLP). This technique can be used as a machine learning tool that allows for an objective evaluation of text-based data without coding bias (Chowdhury, 2003; Hirschberg & Manning, 2015; Nadkarni, Ohno-Machado, & Chapman, 2011). Multiple uses have already demonstrated the efficacy of NLP methods for real-world applications. NLP has been used to mine social media data to deliver new information about disaster-stricken areas into a central repository (Neubig et al., 2011), applied to measuring organizational culture (Pandey & Pandey, 2019), or even utilized to develop emergency plans for unconventional emergencies (Ni et al., 2023). NLP can also extract hazard information, accident causes, and semantic analysis to inform

accident causation theory (Robinson, 2019; Single, Schmidt, & Denecke, 2020; Yan et al., 2023). These combined applications demonstrate how NLP techniques can be used as an effective tool to provide insight into the safety climate of an organization or enhance safety applications despite dynamic conditions.

The current study applied NLP to open-ended questions about the safety climate in an operational naval environment. Rather than allow observer bias to guide the responses, this approach permits the discovery of naturally occurring trends where the sailors themselves can determine the frequency of response for different effective or ineffective safety climate issues—that is, let the sailors themselves determine what is positive or negative about the safety climate. To minimize the influence while collecting the necessary information, participants answered 3 open-ended questions as part of a larger survey effort: 1) sharing their thoughts about safety-related or crew endurance issues; 2) describe the best safety-related behaviors/practices; and, 3) describe the worst safety-related behaviors/practices. Analyses used NLP to determine the topic most represented by the responses, and in turn, the current findings should inform evaluations of naval safety culture by exploring what responses sailors give without the potential influence of observer bias guiding their responses.

2. Methods.

2.1. Data Sample and Questions. Analyses utilized data from the Afloat Safety Climate Assessment Survey (ASCAS; Russell, Russell, & Lei, 2022). This survey tool was developed by the United States Navy to evaluate shipboard organizational and safety climate issues onboard surface force ships. When ship complete the survey, each crew member receives an invitation to participate. The survey landing page provided an overview of the survey's purpose and informed

them that the survey was voluntary and anonymous. Between November 2020 and October 2022, a total of 26,704 individuals were contacted to partake in the survey with 11,917 individuals completing the ASCAS.

For the current analyses, three questions were included in the analyses as the open-ended items. The first question read, “please feel free to share any thoughts that you might have concerning safety-related issues (e.g., ideas to improve workplace safety) and/or Crew Endurance issues (e.g., ideas to improve onboard sleep, exercise, and nutrition) at your current command.” This question provided an opportunity for the individuals to share any of their thoughts related to safety principles aboard ship with minimal guidance. Two subsequent questions attempted to determine whether crew member’s experiences should be considered a best practice or a worst practice. One question asked participants, “please describe the BEST safety-related behaviors/practices that you have witnessed at your current command”. This question captured safety-related factors the individual viewed positively. Conversely, another question asked participants, “please describe the WORST safety-related behaviors/practices that you have witnessed at your current command”. This question captured safety-related factors the individual viewed negatively. All three questions provided a text box for participant input. The open-ended general question had a response rate of 55.77% (N = 6,646), the best safety-related practices question had a response rate of 43.57% (N = 5,192), and the worst safety-related practices question had a response rate of 41.94% (N = 4,998).

2.2. Statistical Analyses. NLP can be evaluated using statistical techniques to determine how the results compare against randomized outcomes (Søgaard et al., 2014). The text responses were first cleaned by eliminating dozens of varieties of “nothing to report”, “none”, etc. Preparing the text for modeling was done using the Gensim python package (tokenization,

bigrams, etc). Non-negative Matrix Factorization (NMF) and Latent Dirichlet Allocation (LDA) were the two topic models tested with the number of topics ranging from four to ten. Using coherence as a metric, NMF at six topics was selected as the optimal choice. The output was then conditioned on various demographics (gender, age, ship department, etc.). The significance of the conditioning variable was measured by a chi-squared test of the contingency table. Overall, the most significant variable was paygrade, which was grouped into four bands E1-E3, E4-E6, E7-E9, and Officer. These techniques produce a *p* value comparable to other types of hypothesis testing, albeit with an important caveat. Specifically, the recommended significance testing should be set to a cutoff of $p = .0025$ to ensure the false positive ratio remains less than or equal to 5% (i.e., typical $p < .05$). Therefore, the current analyses will set the limit of significance tests to $p = .0025$, or 2.5×10^{-3}.

3. Results

3.1. Thoughts on Any Safety-related or Crew Endurance Issue

See Figure 1 for raw scores by topic. When asked about any open-ended thoughts related to safety or crew endurance, NLP analysis demonstrated a significant clustering of topics above randomized chance, $p < .001$ (1.35×10^{-13}). The most commonly discussed topics were clearly personnel and readiness for the higher leadership echelons (officers and chief petty officers), although the trend shifted for more junior personnel (see Figure 2). Non-commissioned officers (E6-E4) reported thoughts equally often for personnel/readiness and sleep/watchstanding, but the most junior personnel (E3-E1) identified personnel and readiness the least frequently of their concerns. Instead, their primary feedback involved aspects related to nutrition and habitability, which would include factors such as the quality of food served aboard ship and the comfort of

their berthing compartment. Additional comments clustered into categories of leadership, exercise/personal time, and equipment/working hours. However, there was a clear shift in the most frequently commented topics among the higher echelons of leadership versus the more junior positions.

3.2. Thoughts on the BEST Safety-related Practices

As depicted in Figure 3, when asked about the best safety-related practices, NLP analysis demonstrated a significant clustering of topics above randomized chance, $p < .001$ (4.90×10^{-19}). The majority (>50%) of personnel across all paygrades reported following procedures as the greatest safety-related strength of their command. Secondary topics varied greatly in their frequency. Senior leadership personnel reported excelling in the execution of special evolutions, which would include things such as major exercises or special events that involve high visibility from leadership further up the chain of command. Conversely, junior personnel were more likely to emphasize proper equipment usage and related issues among the best practices of their command.

3.3. Thoughts on the WORST Safety-related Practices

As depicted in Figure 4, when asked the worst safety-related practices, NLP analysis demonstrated a significant clustering of topics above randomized chance, $p < .001$ (6.61×10^{-4}). No topic reached a majority, as with the best safety-related practices, but the same topic was reported most frequently among all groups as the worst safety-related issue—that is, personnel and readiness. All groups reported that the primary issue involved having enough personnel to execute their assigned responsibilities to complete the mission. Each group also commented

upon sleep/watchstanding issues as among the worst safety-related practices at their command. Taken together, the combination of personnel/readiness and sleep/watchstanding did reach a majority of the topics commented upon by all paygrades, and both topics can be affected by having enough personnel to complete the assignment.

Furthermore, it is noteworthy that training and equipment usage were also reported among the worst safety-related practices, in addition to being a topic on the best safety-related practices. More importantly, junior personnel were more likely to cite training as a best practice, whereas senior personnel were more likely to cite training as a worst practice.

4. Discussion

The current study explored safety concerns within an operational naval environment without inadvertently guiding answers using targeted questions but instead collecting open-ended responses to generic question prompts. Data analyses utilized NLP topic modeling techniques to categorize information repeatedly arising among participant responses. The primary demographic variable involved paygrade, which helped distinguish senior and junior personnel responses. Although there were some consistent topics raised across all personnel, there were also some systemic differences due to position in the leadership hierarchy. Specifically, senior leaders were more likely to identify policy implications of the safety practices, whereas junior personnel focused more upon immediate actions—including things such as habitability of the environment and equipment. These findings thus suggest that while all personnel may inhabit the same living-working space in a naval operating environment, leadership positions impose a different view of the safety priorities than junior positions.

Among the most frequently commented topics, and especially as the worst safety-related practices were concerned, two related issues arose repeatedly: personnel/readiness and sleep/watchstanding. Despite representing distinct concepts, these topics are inherently related as they can feedback upon one another. Insufficient personnel can reduce readiness and impose greater stress on the crew, yet the secondary consequence of insufficient personnel may well be deterioration of sleep. Fewer people must continue to handle the same responsibilities, including watchstanding, and so personnel issues can have a pervasive impact throughout the shipboard environment. Sleep in particular is a critical casualty of this shortfall. Evidence has documented how sailors may not receive the requisite sleep recommended per night for good health (Jameson et al., 2022; Matsangas & Shattuck, 2020; Russell et al., 2021), and how impaired sleep can lead to deteriorating capabilities of naval personnel (Brager et al., 2022; Shattuck & Matsangas, 2016; Skornyakov et al., 2017). The current study merely underscores the importance of these influences as prominent topics weighing heavily on the minds of personnel asked to complete these occupational assignments.

Open-ended questions provided a unique opportunity to explore these types of topics, and as a methodology, can present an underused potential to enhance occupational safety. Surveys and predetermined response options can create immensely valuable tools for measuring the safety climate, yet surveyors may unintentionally bias responses based on how they structure questions. Open-ended responses circumvent this problem by allowing respondents the opportunity to prioritize information without the implicit bias of survey prompts. This technique does carry the problem of categorization and data analysis complications, but with the development of novel machine learning techniques, such as NLP, these tools may allow for more insight to be gained from these applications. Still, some topics simply do not offer themselves up

readily for examination without explicit probes. Safety underreporting is likely the prime example of this possibility (Gilbey, Tani, & Tsui,2016; Probst, Barbaranelli, & Petitta, 2013; Probst, Brubaker, & Barsotti, 2008; Probst & Graso, 2013), although issues related to psychological safety (Edmonson, 1999; 2002; 2018; Edmonson & Lei, 2014; Kahn, 1990) might not come through without structure. There are advantages to multiple methods, but the evidence here draws attention to the role survey structure can implicitly play in biasing or guiding responses.

In summary, the current study explored the safety-related concerns of naval personnel when provided the opportunity to structure their own responses. Open-ended questions were analyzed with machine learning techniques to categorize the responses and make meaningful insights from responses. Participants noted that personnel and sleep, related concepts that affect mission completion, were among the most frequently commented topics—and among the worst safety practices at that. Additionally, the paygrade demographic information differentiated that people in more senior leadership positions will be pay more attention to policy-related issues, whereas junior personnel focus upon the daily operations, mentioning factors such as equipment usage more prominently than senior personnel. Taken together, this evidence suggests that personnel, readiness, and sleep concomitantly contribute to the majority of safety-related concerns during naval operations.

References

Allianz. (2019). Safety and shipping review 2019: An annual review of trends and developments in shipping losses and safety. https://www.agcs.allianz.com/content/dam/onemarketing/agcs/agcs/reports/AGCS-SafetyShipping-Review-2019.pdf (Dec. 24, 2019).

Alruqi, W. M., Hallowell, M. R., & Techera, U. (2018). Safety climate dimensions and their relationship to construction safety performance: A meta-analytic review. Safety science, 109, 165-173.

Antão, P., & Soares, C. G. (2019). Analysis of the influence of human errors on the occurrence of coastal ship accidents in different wave conditions using Bayesian Belief Networks. Accident Analysis & Prevention, 133, 105262.

Bartone, P. T., & Barry, C. L. (2011). Leading for resilience in high risk occupations. Working in high risk environments: Developing sustained resilience, 125-141.

Brager, A., McKeon, A., Russell, D. W., & Markwald, R. R. (2022). Military health and performance optimization: a circadian strategy in response to governmental policies. F1000Research, 11(1262), 1262.

Chowdhury, G. G. (2003). Annual review of information science and technology, 37(1), 51–89.

Christian, M. S., Bradley, J. C., Wallace, J. C., & Burke, M. J. (2009). Workplace safety: a meta-analysis of the roles of person and situation factors. Journal of applied psychology, 94(5), 1103-1127.

Clarke, S. (2006). The relationship between safety climate and safety performance: a meta-analytic review. Journal of occupational health psychology, 11(4), 315-327.

Clarke, S. (2010). An integrative model of safety climate: Linking psychological climate and work attitudes to individual safety outcomes using meta-analysis. Journal of Occupational and Organizational psychology, 83(3), 553-578.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. Administrative science quarterly, 44(2), 350-383.

Edmondson, A. C. (2002). Managing the risk of learning: Psychological safety in work teams (pp. 255-275). Cambridge, MA: Division of Research, Harvard Business School.

Edmondson, A. C. (2018). The Fearless Organization: Creating Psychological Safety in the Workplace for Learning, Innovation, and Growth. Hoboken, NJ: John Wiley & Sons.

Edmondson, A. C., & Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. Annual review of organizational psychology and organizational behavior, 1(1), 23-43.

Gilbey, A., Tani, K., & Tsui, W. H. K. (2016). Outcome knowledge and under-reporting of safety concerns in aviation. Applied Cognitive Psychology, 30(2), 141-151.

Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. Journal of Occupational Health Psychology, 5(3), 347–358.

Hänninen, M. (2014). Bayesian networks for maritime traffic accident prevention: Benefits and challenges. Accident Analysis & Prevention, 73, 305-312.

Harrison, E., Glickman, G. L., Beckerley, S., & Taylor, M. K. (2017). Self-reported sleep during US Navy operations and the impact of deployment-related factors. Military medicine, 182(suppl_1), 189-194.

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, *349*(6245), 261-266.

Jameson, J. T., Markwald, R. R., Kubala, A. G., Roma, P. G., Biggs, A. T., Lai, K., & Russell, D. W. (2022). Sleep deficiency, operational fatigue and the interplay of compromising factors: Analysis to aid in fatigue management. Journal of Sleep Research.

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. Academy of management journal, 33(4), 692-724.

Krentz, M. J., Li, G., & Baker, S. P. (1997). At work and play in a hazardous environment: injuries aboard a deployed US Navy aircraft carrier. Aviation Space and Environmental Medicine, 68(1), 51-55.

Leigh, J. P., Marcin, J. P., & Miller, T. R. (2004). An estimate of the U.S. Government's undercount of nonfatal occupational injuries. Journal of Occupational and Environmental Medicine, 46, 10 –18.

Lowery, J. T., Borgerding, J. A., Zhen, B., Glazner, J. E., Bondy, J., & Kreiss, K. (1998). Risk factors for injury among construction workers at Denver International Airport. American Journal of Industrial Medicine, 34, 113–120.

Matsangas, P., & Shattuck, N. L. (2017, September). Exploring sleep-related habitability issues in berthing spaces on US Navy ships. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 61, No. 1, pp. 450-454). Sage CA: Los Angeles, CA: SAGE Publications.

Matsangas, P., & Shattuck, N. L. (2020). Sleep quality, occupational factors, and psychomotor vigilance performance in the US Navy sailors. *Sleep*, *43*(12), zsaa118.

Matsangas, P., & Shattuck, N. L. (2021). Habitability in Berthing Compartments and Well-Being of Sailors Working on U.S. Navy Surface Ships. Human Factors, 63(3), 462-473.

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, *18*(5), 544-551.

Neubig, G., Matsubayashi, Y., Hagiwara, M., & Murakami, K. (2011, November). Safety Information Mining—What can NLP do in a disaster—. In *Proceedings of 5th International Joint Conference on Natural Language Processing* (pp. 965-973).

Neal, A., & Griffin, M. A. (2002). Safety climate and safety behaviour. Australian journal of management, 27(1_suppl), 67-75.

Neal, A., & Griffin, M. A. (2004). Safety climate and safety at work. In J. Barling & M. R. Frone (Eds.), The psychology of workplace safety (pp. 15–34). Washington, DC: American Psychological Association.

Neal, A., Griffin, M. A., & Hart, P. M. (2000). The impact of organizational climate on safety climate and individual behavior. Safety science, 34(1-3), 99-109.

Ni, W., Shen, Q., Liu, T., Zeng, Q., & Xu, L. (2023). Generating textual emergency plans for unconventional emergencies—A natural language processing approach. *Safety Science*, *160*, 106047.

Pandey, S., & Pandey, S. K. (2019). Applying natural language processing capabilities in computerized textual analysis to measure organizational culture. *Organizational Research Methods*, *22*(3), 765-797.

Probst, T. M. (2015). Organizational safety climate and supervisor safety enforcement: Multilevel explorations of the causes of accident underreporting. Journal of applied psychology, 100(6), 1899-1907.

Probst, T. M., Barbaranelli, C., & Petitta, L. (2013). The relationship between job insecurity and accident under-reporting: A test in two countries. Work & Stress, 27(4), 383-402.

Probst, T. M., Brubaker, T. L., & Barsotti, A. (2008). Organizational injury rate underreporting: the moderating effect of organizational safety climate. Journal of Applied Psychology, 93(5), 1147-1154.

Probst, T. M., & Estrada, A. X. (2010). Accident under-reporting among employees: Testing the moderating influence of psychological safety climate and supervisor enforcement of safety practices. Accident analysis & prevention, 42(5), 1438-1444.

Probst, T. M., & Graso, M. (2013). Pressure to produce = pressure to reduce accident reporting? Accident Analysis and Prevention, 59, 580 –587.

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in web questionnaires. *Developments in applied statistics*, *19*(1), 159-177.

Robinson, S. D. (2019). Temporal topic modeling applied to aviation safety reports: A subject matter expert review. *Safety science*, *116*, 275-286.

Roethlisberger, F.J., & Dickson, W.J. (1967). Management and the worker. Cambridge, MA: Harvard University Press. (Original work published 1939).

Russell, D. W., Markwald, R. R., & Jameson, J. T. (2021). Self-reported sleep and sleep deficiency: Results from a large initiative of sailors attached to US Navy warships. Journal of Sleep Research, 30(6), e13397.

Russell, D. W., Russell, C. A., & Lei, Z. (2022). Development and testing of a tool to measure the organizational safety climate aboard US Navy ships. Journal of safety research, 80, 293-301.

Schmidt, J., Schmorrow, D., & Figlock, R. (2000, July). Human factors analysis of naval aviation maintenance related mishaps. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 44, No. 22, pp. 775-778). Sage CA: Los Angeles, CA: SAGE Publications.

Shattuck, N. L., & Matsangas, P. (2016). Operational assessment of the 5-h on/10-h off watchstanding schedule on a US Navy ship: sleep patterns, mood and psychomotor vigilance performance of crewmembers in the nuclear reactor department. *Ergonomics*, *59*(5), 657-664.

Shattuck, N. L., & Matsangas, P. (2017). Sunlight exposure, work hours, caffeine consumption, and sleep duration in the naval environment. Aerospace Medicine and Human Performance, 88(6), 579-585.

Shattuck, N. L., & Matsangas, P. (2022). Who sleeps more and who works longer in the US Navy: Officers or enlisted personnel?. Sleep Health, 8(4), 387-390.

Single, J. I., Schmidt, J., & Denecke, J. (2020). Knowledge acquisition from chemical accident databases using an ontology-based method and natural language processing. *Safety Science*, *129*, 104747.

Skornyakov, E., Shattuck, N. L., Winser, M. A., Matsangas, P., Sparrow, A. R., Layton, M. E., ... & Van Dongen, H. P. (2017). Sleep and performance in simulated Navy watch schedules. *Accident Analysis & Prevention*, *99*, 422-427.

Søgaard, A., Johannsen, A., Plank, B., Hovy, D., & Alonso, H. M. (2014, June). What's in a p-value in NLP?. In Proceedings of the eighteenth conference on computational natural language learning (pp. 1-10).

U.S. Navy, 2017a. Comprehensive review of surface force incidents. Norfolk, VA. http://s3.amazonaws.com/chinfo/comprehensive+review_final.pdf (Dec. 12, 2019).

Wickström, G., & Bendix, T. (2000). The" Hawthorne effect"—what did the original Hawthorne studies actually show?. *Scandinavian journal of work, environment & health*, 363-367.

Yan, K., Wang, Y., Jia, L., Wang, W., Liu, S., & Geng, Y. (2023). A content-aware corpus-based model for analysis of marine accidents. *Accident Analysis & Prevention*, *184*, 106991.

Zhang, G., & Thai, V. V. (2016). Expert elicitation and Bayesian Network modeling for shipping accidents: A literature review. Safety science, 87, 53-62.

Zohar, D. (1980). Safety climate in industrial organizations: theoretical and applied implications. Journal of applied psychology, 65(1), 96.

Zohar, D. (2010). Thirty years of safety climate research: Reflections and future directions. Accident analysis & prevention, 42(5), 1517-1522.

Zohar, D. (2011). Safety climate: Conceptual and measurement issues. In D. Hofmann and L. Tetrick (Eds.), The Handbook of Occupational Health Psychology (pp. 141-164). Washington, DC: American Psychological Association.

Zohar, D., & Luria, G. (2004). Climate as a social-cognitive construction of supervisory safety practices: scripts as proxy of behavior patterns. Journal of applied psychology, 89(2), 322-333.

Table 1. Overview of categories derived from topic modeling.

Open-Ended Question	Category	Topics
Thoughts on safety-related issues (e.g., ideas to improve workplace safety) and/or Crew Endurance issues (e.g., ideas to improve onboard sleep, exercise, and nutrition)	Personnel & Readiness	Ship, Sailor, Navy, Man, Need, Training
	Sleep & Watchstanding	Sleep, Watch, Hour, Day, Night, Stand
	Nutrition & Habitability	Food, Need, Option, Healthy, Portion, Rack
	Leadership	Command, Good, Chain, Care, Job, PT
	Exercise & Personal Time	Time, Exercise, Personal, People, Schedule, PT
	Equipment & Working Hours	Work, Day, Hour, Equipment, Complete, Need
Describe the BEST safety-related behaviors/practices that you have witnessed at your current command	Following Procedures	Good, Practice, Follow, Procedure, Time, Sailor
	Protective Equipment	PPE, Wear, Proper, Use, Work, Maintenance
	Special Evolutions	Evolution, Debrief, Major, Conduct, Prior, Special
	Training	Training, Conduct, Consistent, Lot, Stand, Require
	Planning	Brief, ORM, Prior, Pre, Conduct, Plan
	Equipment Usage	Aloft, Harness, Work, Program, Wear, Use
Describe the WORST safety-related behaviors/practices that you have witnessed at your current command?	Personnel & Readiness	People, Ship, Care, Issue, Time, Ladder
	Sleep & Watchstanding	Work, Time, Sleep, Day, Watch, Hour
	Training	Sailor, Bad, Junior, Ship, Practice, Know
	Equipment Usage	PPE, Wear, Use, Proper, Harness, Paint
	Safety Reporting	Report, Issue, Failure, Moment, Practice, Unsafe
	Falsifying Reports	NTR, Procedure, Follow, Deck, Disregard, Gun

OpenEnded: Thoughts on safety-related issues and Crew Endurance issues | 27.0% response rate | -12.87 log(p-value)

topics:	ship sailor navy man need training	sleep watch hour day night stand	food need option healthy portion rack	command good chain care pt job	time exercise personal people schedule pt	work day hour equipment complete need
Paygrade						
E1-E3	12.79	17.44	24.81	13.95	15.12	15.89
E4-E6	20.64	20.57	15.71	15.26	13.61	14.21
E7-E9	33.94	13.87	10.58	16.06	17.52	8.03
Officer / CWO	32.14	20.48	11.67	13.57	14.76	7.38

BEST safety-related behaviors witnessed at your current command | 20.0% response rate | -18.31 log(p-value)

topics:	good practice follow procedure time sailor	ppe wear proper use work maintenance	evolution debrief major conduct prior special	training conduct consistent lot stand require	brief orm prior pre conduct plan	aloft harness work program wear use
Paygrade						
E1-E3	50.97	19.42	5.83	10.68	6.80	6.31
E4-E6	57.67	11.67	5.78	10.56	6.00	8.33
E7-E9	52.21	7.96	15.49	5.31	12.83	6.20
Officer / CWO	53.39	4.88	19.51	7.05	11.38	3.79

WORST safety-related behaviors witnessed at your current command? | 20.0% response rate | -3.18 log(p-value)

topics:	people ship care issue time ladder	work time sleep day watch hour	sailor bad junior ship practice know	ppe wear use proper harness paint	report issue failure moment practice unsafe	ntr procedure follow deck disregard gun
Paygrade						
E1-E3	35.06	26.62	14.29	22.08	1.95	0.00
E4-E6	33.41	28.96	20.87	13.68	2.62	0.46
E7-E9	30.61	26.53	23.98	11.73	4.08	3.06
Officer / CWO	29.18	27.87	22.62	12.79	4.26	3.28

Figure 1. Raw values for the responses by topic.

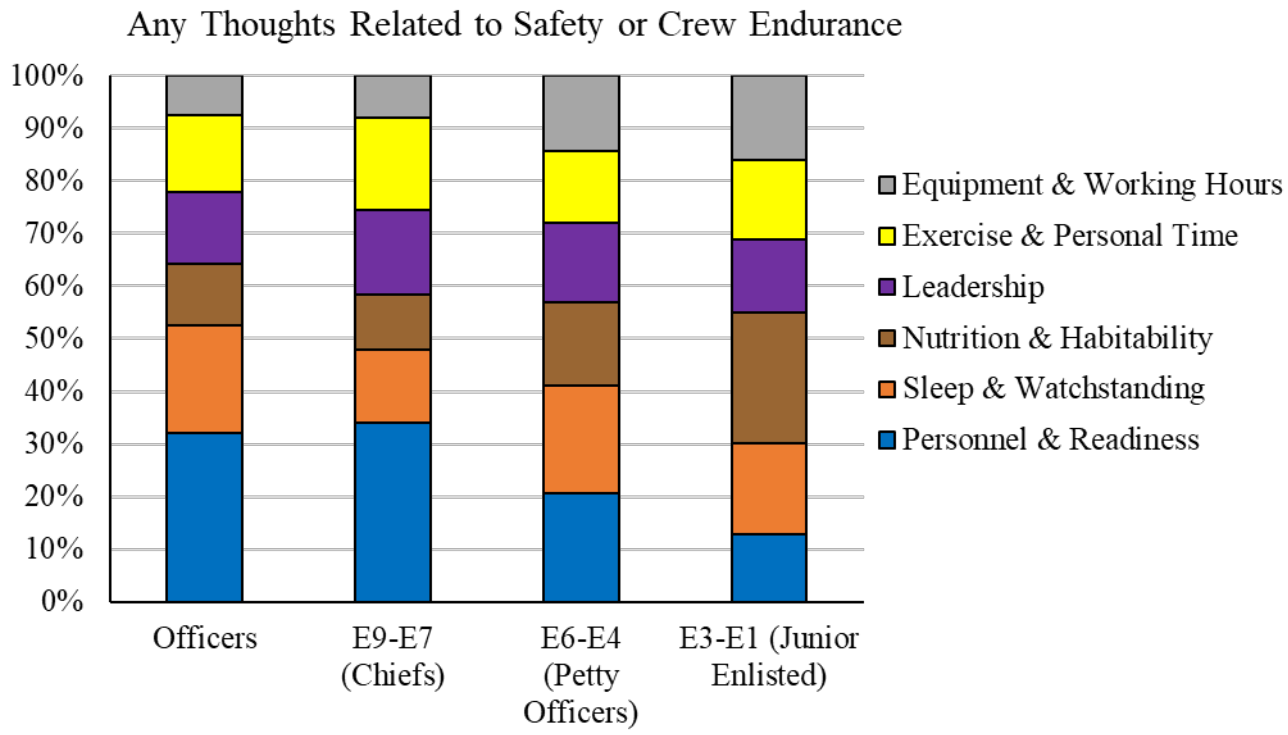

Figure 2. Topic response frequency when asked to give feedback about any safety-related or crew endurance topic.

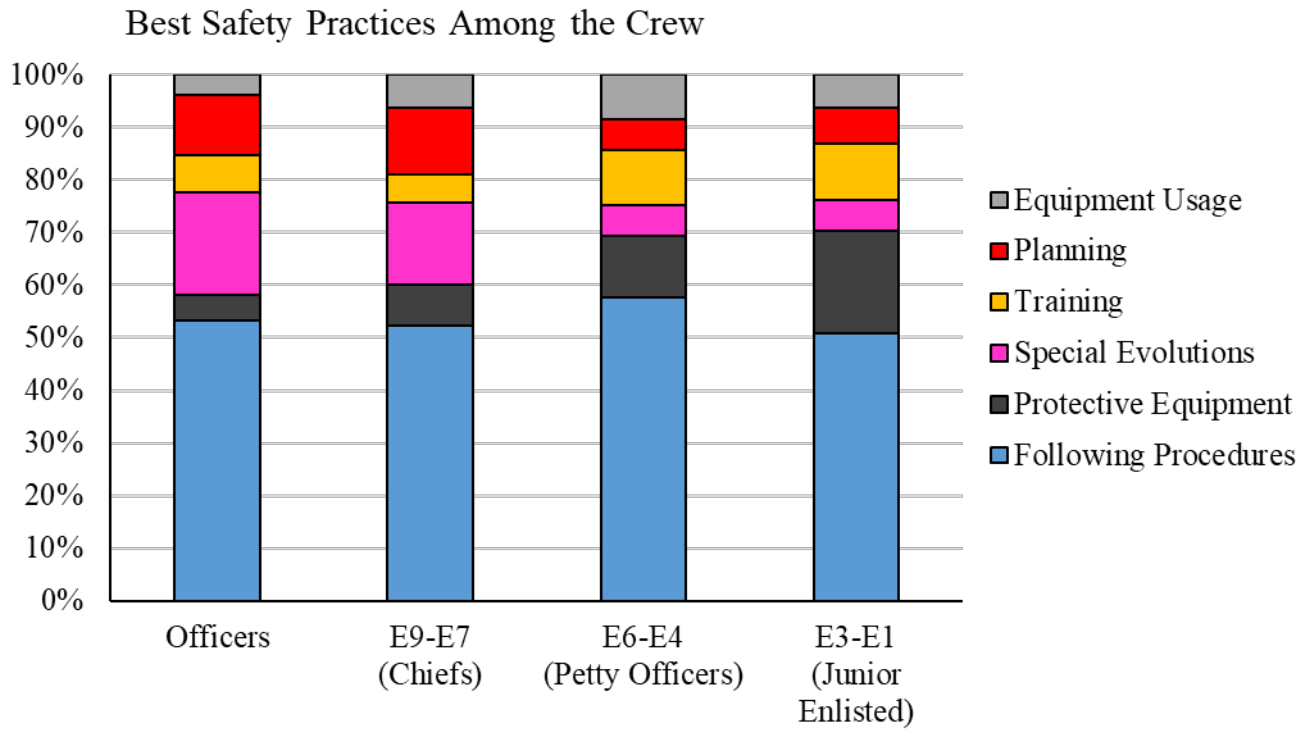

Figure 3. Topic response frequency when asked to give feedback about the best safety-related practices aboard ship.

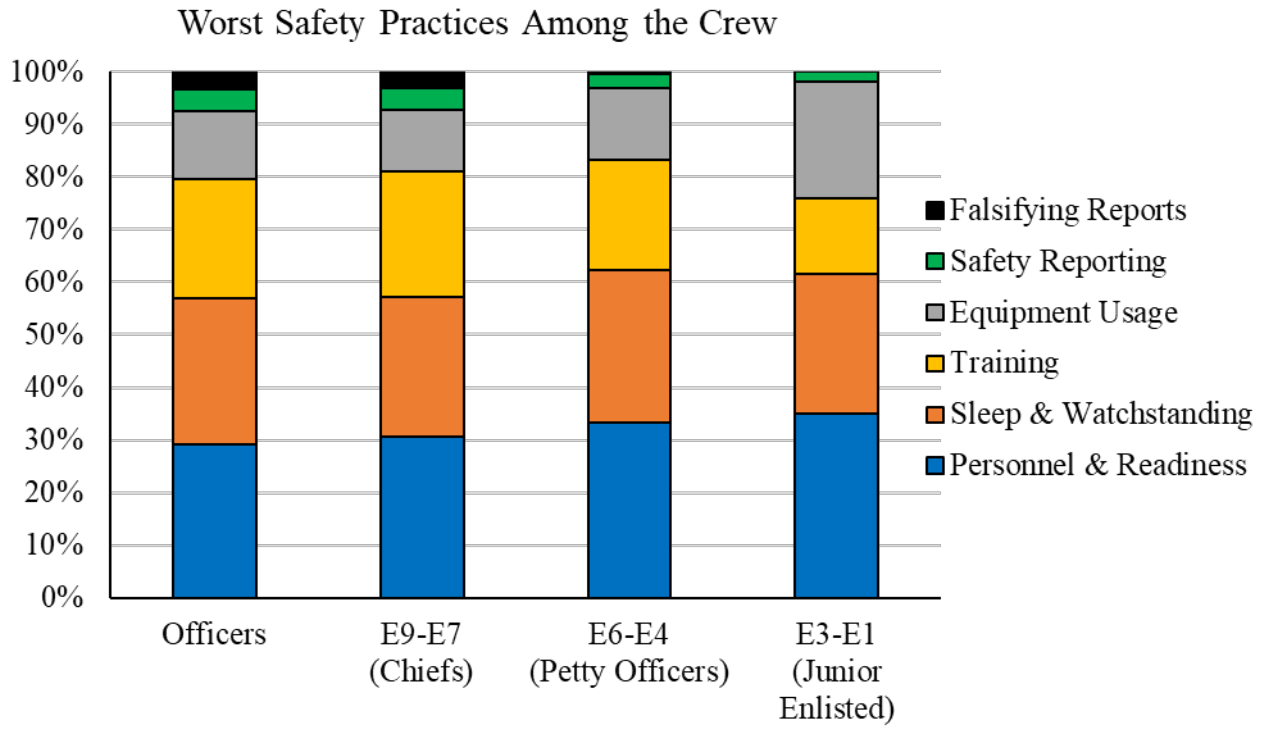

Figure 4. Topic response frequency when asked to give feedback about the worst safety-related practices aboard ship.