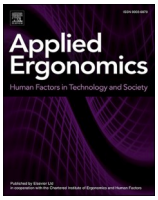
Applied Ergonomics 117 (2024) 104225

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo



Advancing a U.S. navy shipboard infrastructure for sleep monitoring with wearable technology

Andrew G. Kubala [a,b,1], Peter G. Roma [a,b,1], Jason T. Jameson [a,b], Pinata H. Sessoms [a], Evan D. Chinoy [a], Luis R. Rosado [a,b], Trevor B. Viboch [a,b], Brandon J. Schrom [a,b], Hedaya N. Rizeq [a,b], Prayag S. Gordy [a,b], LT Dale A. Hirsch [c], LCDR Adam T. Biggs [d], Dale W. Russell [e,f], Rachel R. Markwald [a,*,1]

[a] Warfighter Performance Department, Operational Readiness and Health Directorate, Naval Health Research Center, San Diego, CA, USA
[b] Military and Veterans Health Solutions, Leidos Inc., San Diego, CA, USA
[c] Navy Leadership and Ethics Center, Newport, RI, USA
[d] Psychological Health and Resilience Department, Military Population Health Directorate, Naval Health Research Center, San Diego, CA, USA
[e] Commander Naval Surface Force, U.S. Pacific Fleet, San Diego, CA, USA
[f] Department of Psychiatry, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

ARTICLE INFO

Keywords:
Wearables
Sleep
Fatigue risk management

ABSTRACT

Development of fatigue management solutions is critical to U.S. Navy populations. This study explored the operational feasibility and acceptability of commercial wearable devices (Oura Ring and ReadiBand) in a warship environment with 845 Sailors across five ship cohorts during at-sea operations ranging from 10 to 31 days. Participants were required to wear both devices and check-in daily with research staff. Both devices functioned as designed in the environment and reliably collected sleep-wake data. Over 10,000 person-days at-sea, overall prevalence of Oura and ReadiBand use was 69% and 71%, respectively. Individual use rates were 71 ± 38% of days underway for Oura and 59 ± 34% for ReadiBand. Analysis of individual factors showed increasing device use and less device interference with age, and more men than women found the devices comfortable. This study provides initial support that commercial wearables can contribute to infrastructures for operational fatigue management in naval environments.

1. Introduction

Sleep-related fatigue degrades performance (Grandou et al., 2019; Charest and Grandner, 2022), increases the risk of errors (Williamson et al., 2011), and is a key contributing factor to negative operational outcomes in high reliability organizations (HROs) such as those in the aviation (Rosekind et al., 1994; Bendak and Rashid, 2020), medical (Kancherla et al., 2020; Barger et al., 2018), and commercial transportation industries (Connor et al., 2001; Taylor and Dorn, 2006; Stern et al., 2019). As such, HROs must surveille and mitigate sleep-related fatigue risks for workers "operating in stressful situations involving complex environments, high degrees of uncertainty, time pressure, and severe consequences for mistakes" (Baumann et al., 2011). Successful HROs have implemented evidence-based safety practices and have established work climates that uphold responsibility and accountability to reduce the incidence of operational safety issues and may adopt operational risk management (ORM) processes such as fatigue management systems (Roberts, 1990; Manuele, 2013) to help identify and minimize risk.

A popular fatigue-related ORM process includes limiting work hours to provide the opportunity for sufficient sleep and subsequently reduce fatigue and improve alertness (Caldwell et al., 2008; Peets and Ayas, 2012). However, limiting work hours in the military context is not always a practical intervention due to operational realities; accordingly, improved monitoring approaches (e.g., wearable-based personal status monitoring solutions) hold promise to help leaders account for fatigue-related risk. But the process of developing effective, operationally-embedded monitoring solutions is complex and challenging, thus sleep-related fatigue and human error continue to serve as a source of mishaps and near-misses in civilian maritime organizations (Sánchez-Beaskoetxea et al., 2021; Islam et al., 2020) and operational

* Corresponding author. Warfighter Performance Department Naval Health Research Center, 140 Sylvester Road, San Diego, CA, 92106, USA.
E-mail address: rachel.r.markwald.civ@health.mil (R.R. Markwald).
[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.apergo.2024.104225
Received 31 July 2023; Received in revised form 20 December 2023; Accepted 3 January 2024
Available online 14 January 2024

Abbreviations

HRO	High Reliability Organization
ORM	Operational Risk Management
USN:	United States Navy
SURFOR	Surfaces Forces
GAO	US Government Accountability Office
PSG	Polysomnography
PSM	Personal Status Monitors
PPG	Photoplethysmography
CREW	Command Readiness, Endurance, and Watchstanding Program
DDG	Arleigh Burke-class destroyer ship
LCS	Independence-class littoral combat ship
LHD	Wasp-class landing helicopter dock ship
OR	Oura Ring
RB	ReadiBand
TST	Total Sleep Time
WASO	Wake After Sleep Onset
EEG	Electroencephalography
TIB	Time in Bed
SOL:	Sleep Onset Latency

naval forces (Bierly and Spender, 1995; Gould et al., 2006).

Maritime operations, in general, present unique occupational (e.g., shiftwork, long work hours) and environmental (e.g., austere work/habitability settings, dynamic weather and sea state conditions) stressors that threaten operational safety. For instance, fatigue-related human error and long work hours were contributing factors to the Exxon Valdez tanker mishap that resulted in approximately 11 million gallons of crude oil spilled into the Prince William Sound (National Transportation Safety Board, 1990). In addition to facing routine maritime operational stressors, naval personnel onboard warships are faced with maintaining a continuous state of combat readiness over extended durations while often experiencing undermanned crews, limited healthful meal and exercise options, uncomfortable sleeping quarters, and limited connection to social support networks back home (Matsangas and Shattuck, 2021; Russell et al., 2021; Gottschall and Guérin, 2021). These considerations have made the implementation of consistent, technology-assisted fatigue management processes a difficult task in naval operations.

For the United States Navy (USN), Surface Force personnel (SURFOR, which comprises all warships other than aircraft carriers and submarines) routinely lack adequate sleep (Russell et al., 2021; Jameson et al., 2022; Matsangas and Shattuck, 2020). The risks associated with such sleep deficiencies came to the forefront in 2017 following two major mishaps that resulted in the deaths of 17 sailors and hundreds of millions of dollars in damage. The United States Government Accountability Office (GAO) concluded that fatigue was a contributing factor in both mishaps, and recommended that the USN take action to enable methods and policies to collect timely and quality sleep and fatigue-related data that are available to Commanders for decision-making at sea (U.S. Government Accountability Office, 2021). In response, SURFOR undertook numerous efforts to address sleep deficiency and reduce or eliminate fatigue-related risk, such as assessing the ability to leverage commercial wearable personal status monitors (PSMs) to monitor sleep and implement software programs to effectively predict fatigue-related risks (Eckstein, 2022a, 2022b). Such a capability would significantly advance the USN's fatigue management practices and help enhance the fleet's operational safety.

Schedules and subsequent sleep opportunities often change during naval operations, and sleep monitoring technologies must be accurate, reliable, and require minimal burden on busy sailors if they are to be viable solutions. While polysomnography (PSG) is considered the gold standard for sleep assessment in laboratory and clinical settings, it is not a viable field-based solution for long-term sleep monitoring due to the complex physical setup and time burden to process and interpret the data. Similarly, the popular field-standard method of research-grade actigraphy is also impractical as this process requires manual uploading of data from a wrist-worn device and specialized software for interpretation by trained technicians. Thus, the desired capability for monitoring sleep and predicting fatigue risk requires on-demand data technology that is easy to use, automatic, and requires no technical expertise (Chinoy et al., 2023a).

In recent years the commercial PSM market has witnessed an explosion of devices that are capable of continuously measuring physiology and estimating sleep (Ferreira et al., 2021). Commercial PSMs estimate sleep parameters using activity data derived from tri-axial accelerometers and, in some cases, blood volume dynamics (e.g., heart rate and respiration rate) derived from photoplethysmography (PPG) sensors (Rentz et al., 2021). Using a wireless connection (e.g., Bluetooth), PSMs then communicate with software applications on smart devices where data processing occurs automatically, either on the device or within a commercial cloud environment, then relays data feedback to the user. Research from the author group assessing the accuracy of PSMs has found many perform as well as or better than actigraphy for sleep and wake classifications (Chinoy et al., 2021, 2022; Stone et al., 2020). Although PSMs could provide automated and reliable on-demand data for downstream fatigue management systems (Reifman et al., 2023), the technical feasibility and acceptability of using PSMs in austere warship environments (i.e., turbulent sea states, internal environmental conditions, signal interference, workload) has not been empirically evaluated.

The primary aim of this study was to provide an initial assessment of technical feasibility and operational acceptability (i.e., device use, user experience) of PSMs onboard warships while at-sea. The secondary aim was to explore the impact of device and demographic characteristics on operational acceptability.

2. Methods

2.1. Recruitment and study design

This study recruited active duty USN personnel assigned to SURFOR warships as part of the Command Readiness, Endurance, and Watchstanding (CREW) research program. Warships were identified by SURFOR to participate in the study based on their at-sea availability schedules. The crews of each ship attended a recruitment session where they were provided a study overview and informed of their rights. All ship crewmembers were eligible to participate. Those interested in participating provided written informed consent. The study protocol was approved by the Naval Health Research Center Institutional Review Board (Protocol # NHRC.2021.0003) in compliance with all applicable Federal regulations governing the protection of human subjects.

The current investigation used an observational cohort design in which participants were asked to: (1) complete a baseline questionnaire that obtained demographic, sleep quality, and health-related behaviors data; (2) wear two distinct PSMs from the time of enrollment to the end of their at-sea mission (Oura Ring [Oura Health Ltd., Oulu, Finland] and ReadiBand [Fatigue Science, Vancouver, Canada]); (3) check-in daily with onboard research staff; and (4) complete a final questionnaire that assessed PSM user experience. Data collection occurred across three ship-classes that varied in size and mission, with at-sea periods ranging from 10 to 31 days.

2.2. Sample selection

Four warship cohorts provided the opportunity to collect data during five underway periods. The cohorts included two Arleigh Burke-class destroyer ships (DDG; hereafter referred to as DDG-A and DDG-B; 12

and 31 days underway, respectively), one Independence-class littoral combat ship (LCS-A; 10 days underway), and two separate cohorts onboard one Wasp-class landing helicopter dock (LHD) ship (LHD-A and LHD-B; 13 and 16 days underway, respectively).

A total of 853 personnel consented to participate. Participants were excluded from analysis if they did not go underway following enrollment, did not complete the baseline questionnaire prior to going underway, voluntarily withdrew from the study prior to going underway, or experienced a technical error(s) resulting in no data or lost data from a PSM. The final sample for analysis consisted of 845 personnel.

2.3. Measures

2.3.1. Commercial off-the-shelf PSMs

Participants were fitted with a Generation 2 Oura Ring (OR) and a ReadiBand (RB) version 4 that each calculate measures of sleep-wake behavior using proprietary algorithms. The OR is a finger-worn, water-resistant, multi-sensor wearable device that measures physiology (e.g., heart rate) via PPG, body temperature via negative temperature coefficient sensor, and movement via tri-axial accelerometer. In laboratory validation studies, the OR has shown good agreement with PSG in whole night estimation of total sleep time (TST) and wake after sleep onset (WASO) (de Zambotti et al., 2019; Chee et al., 2021). Additionally, in free-living environments, the OR performed similarly, or better, than actigraphy compared with ambulatory electroencephalography (EEG) (Chinoy et al., 2022). Participants were fitted with an OR on their preferred hand and finger.

The RB is a wrist-worn device that detects movement via triaxial accelerometry but does not have additional physiological monitoring capabilities. The RB has also been found to perform similarly, or better, than research-grade actigraphy when both are compared to PSG in laboratory and home environments (Chinoy et al., 2021, 2022). The RB was fitted on the same wrist as the OR except when participants expressed concerns regarding work needs or comfort. This flexibility was intended to reduce participant burden and increase participation throughout the ships. Participants were instructed to wear the PSMs whenever possible and to put them back on as soon as possible if ever removed at any point during waking hours (i.e., for operational safety, security, device power re-charging, or personal needs).

The OR was paired to the participant's personal smartphone (53%) or a study-provided smartphone (47%; Samsung A51; Samsung, Seoul, South Korea). RBs were paired to an iPad (Apple Inc., Cupertino, CA) application. Research staff maintained the iPad and A51 devices onboard. Both devices were connected via Bluetooth to their respective applications, but the OR smartphone application did not require cellular connectivity to synchronize. Due to this capability, participants could view their data after their daily synchronization. In contrast, the RB devices required cellular connectivity to synchronize, but contained enough memory that the data were retained until cellular connectivity was possible after the underway collection period. Thus, the RB devices provided no data to the users while at sea. The ORs required re-charging at regular intervals (full charge every 3–4 days, or daily top-off during convenient periods), whereas the RB devices were pre-charged with sufficient power to operate for about 30 days and thus required no power maintenance.

Both the OR and RB devices provide sleep summary measures of time in bed (TIB: the total duration of a recorded sleep episode from bedtime to waketime), total sleep time (TST: total duration of sleep across the TIB interval), sleep efficiency (SE: percentage of TST divided by TIB), sleep onset latency (SOL: time until initially falling asleep after bedtime), and wake after sleep onset (WASO: time awake in each sleep episode occurring after first falling asleep).

During each underway data collection, research staff remained onboard for daily check-ins and to provide technical assistance, except for the DDG-2 cohort, in which a servicemember served as an onsite liaison to handle administrative matters (e.g., study withdrawals) but was unable to support daily check-ins. Study closeout began 2–3 days before returning to port, where study staff ensured a final data upload(s), collected the RB devices, and administered a final questionnaire (described below).

2.3.2. Demographics and PSM user experience

Prior to the underway periods, participants completed a baseline questionnaire and reported standard US Department of Defense demographics of gender (female, male), age (years), race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiracial, or Unknown), and current military paygrade (E1-E3 [junior enlisted], E4-E6 [petty officer], E7-E9 [senior enlisted], Officer/Chief Warrant Officer). Upon study closeout, participants completed a user experience questionnaire to rate both PSMs. The questionnaire operated on a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree) and included the following items: *Wearing the* [device name] *interfered with other activities*; *The* [device name] *was comfortable*; and *I think that I would like to use the* [device name] *frequently* (adapted from the System Usability Scale (Lewis, 2018)).

2.4. Data analysis

Technical feasibility was determined by the rate of nominal device functioning, and the production of orderly sleep data in working devices. Acceptability was defined primarily as device use, as determined per day by the presence of any data indicating active wear/use (i.e., activity data) recorded within each 24-hr period. Acceptability was also defined using self-reported user experience. Device use prevalence is expressed as the percent of all person-days in which the devices were used, both across the entire sample and for each ship cohort. Device use prevalence over time is expressed as the percent of personnel who used the device(s) each day.

Device use at the individual level is expressed as the percentage of days underway in which each participant used their device(s). Device use at the individual level was also categorized in ranges: Non-User groups were issued devices but never used them, the Low Use group used the device(s) between 1 and 25% of days at-sea, the Mid-Low group used the device(s) between 25 and 50% of days, the Mid-High group used the device(s) between 50 and 75% of days, and the High group used the devices between 75 and 100% of days. User experience ratings are presented by PSM device via mean ± standard deviation (SD) as well prevalence by response item.

To examine predictors of individual PSM use, we used a linear mixed effect model with fixed factors of device (OR, RB), age (continuous), gender, and device x gender and device x age interactions. Ship cohort and race were included as cluster variables, with random intercepts for each. To examine predictors of PSM acceptability, we applied generalized mixed models for ordinal logistic regression to the 5-point scale responses of each of the user experience questions (interference, comfort, and want to use). Each analysis included fixed factors of device, age, gender, and device x gender and device x age, interactions with cluster variables of ship cohort and race as random intercepts. Statistical significance was set at $\alpha = .05$. Data processing and analyses were performed in the R statistical programming language (R Foundation for Statistical Computing, Vienna, Austria) and jamovi 2.4 (The jamovi project, Sydney, Australia).

3. Results

3.1. Sample characteristics

The study sample was consistent with active duty USN population demographics in gender (81% male), race (54% White), ethnicity (25% Hispanic or Latino), age (18–25 = 44%, 26–35 = 37%, ≥36 = 19%), and proportions of enlisted personnel and officers (83% and 16%,

respectively; see Table 1). (U.S. Department of Defense, 2023) Overall, 46% of eligible personnel consented to participate (range 27–87% per ship cohort; see Supplementary Table 1).

3.2. PSM technical feasibility

Both the OR and RB devices performed as designed within the warship environments, with no systematic technical or environmental impediments to nominal function observed. Technical malfunction or failure resulting in no data capture or data transfer occurred in 10 OR devices (1.2%) and 19 RB devices (2.2%). Mean ± SD total TIB per day was 413 ± 106 min in the OR and 392 ± 138 min in the RB; mean TST was 348 ± 89 min in the OR and 386 ± 127 min in the RB; mean sleep efficiency was 82 ± 16% in the OR and 85 ± 10 min in the RB; mean SOL was 11 ± 12 min in the OR and 21 ± 31 min in the RB; and mean WASO was 54 ± 38 min in the OR and 30 ± 36 min in the RB (see Supplementary Table 2).

3.3. PSM acceptability: device use

Overall, the OR was used during 69% of the 10,555 applicable person-days underway and the RB was used on 71% of 10,424 person-days. Prevalence of device use varied by cohort, with OR ranging from 43 to 83%, and RB ranging from 52 to 84% (see Supplementary Table 3). As seen in Fig. 1, aggregate device use prevalence decreased over time, with an average 1% loss of personnel for each day underway.

At the individual level, personnel used the OR for 71 ± 38% of days underway, and the RB for 60 ± 33% of days. Individual device use rates also varied by cohort, with mean ± SD use of the OR ranging from 42 ± 45% to 80 ± 33%, and for the RB ranging from 51 ± 48% to 65 ± 34% (see Supplementary Table 4). As shown in Fig. 2, the distribution of individual PSM use varied by device. 14% of personnel did not use the OR, and 14% did not use the RB; another 12% used the OR less than half of the days at-sea, and 14% used the RB less than half of the days at-sea. Most personnel, 68%, used the OR for at least 75% of days underway, in contrast to only 44% of personnel for the RB. Supplementary Figure S1 presents individual use patterns across days for all personnel.

Analysis of individual device use revealed significant main effects of device and age ($Fs > 24.00$, $ps < .001$) and a device × age interaction ($F(1,1599) = 5.27$, $p < .05$; all other terms $Fs < 1.10$, $ps > .25$). Likelihood ratio tests of random effects supported significant model improvement by inclusion of ship cohort (X^2 (1) = 33.14, $p < .001$), but not race (X^2 (1) = 3.50, $p = .06$). As shown in Fig. 3, OR use was higher than RB, and the use of both devices increased as a function of age, with a stronger age-use relationship in the OR than RB.

Table 1
Participant demographics.

	Demographic	*N*	%
Gender	Female	161	19.03
	Male	685	80.97
Age	18–25	364	43.96
	26–35	308	37.20
	36+	156	18.84
Race	American Indian or Alaska Native	18	2.13
	Asian	103	12.19
	Black or African American	160	18.93
	Native Hawaiian or Other Pacific Islander	8	0.95
	White	459	54.32
	Multi-Racial	40	4.73
	Unknown	57	6.75
Hispanic or Latino	No	470	74.48
	Yes	161	25.52
Paygrade	E1-E3	190	22.54
	E4-E6	416	49.35
	E7-E9	104	12.34
	Officer	133	15.78
All		845	

3.4. PSM acceptability: user experience

Mean ratings of device interference were 2.23 ± 1.14 for the OR and 2.30 ± 1.23 for the RB, ratings of device comfort were 3.65 ± 1.16 for the OR and 3.53 ± 1.26 for the RB, and ratings of wanting to use the device were 4.21 ± 0.97 for the OR and 2.97 ± 1.30 for the RB (see Supplementary Table 5). Fig. 4 provides the distribution of user experience ratings between device type. Overall, device interference and device comfort followed similar distributions between devices, with approximately 58–62% of participants rating they disagreed or strongly disagreed that the devices interfered with daily activities. Additionally, 54–58% of participants agreed or strongly agreed that the devices were comfortable. More participants (78%) reported agreeing or strongly agreeing to wanting to use the OR more frequently in the future compared to 33% wanting to use the RB.

Analysis of individual predictors of device interference ratings revealed a significant main effect of age (X^2 (1) = 6.54, $p < .05$), whereby older participants were more likely to strongly disagree with the statement that wearing the devices interfered with other activities (all other terms $X^2s < 1.55$, $ps > .20$; see Supplementary Figure 2). Likelihood ratio tests of random effects supported significant model improvement by inclusion of ship cohort and race ($X^2s > 15.80$, $ps < .001$). Analysis of individual predictors of device comfort ratings revealed a significant main effect of gender (X^2 (1) = 5.99, $p < .05$), whereby men were more likely than women to rate the devices as comfortable (all other terms $X^2s < 1.00$, $ps > .25$). Likelihood ratio tests of random effects supported significant model improvement by inclusion of ship cohort and race ($X^2s > 10.75$, $ps < .01$). Analysis of individual predictors of wanting to use the device revealed a significant main effect of device (X^2 (1) = 5.99, $p < .05$), whereby personnel were more likely to agree or strongly agree with wanting to use the OR than the RB (all other terms $X^2s < 3.00$, $ps > .08$). Likelihood ratio tests of random effects supported significant model improvement by inclusion of ship cohort and race ($X^2s > 8.80$, $ps < .01$).

4. Discussion

Inadequate sleep and resulting fatigue increase the risk of human errors, especially in operational environments such as those routinely experienced by military personnel. An active fatigue management system, within a larger safety management system, holds potential in helping prevent fatigue-related mishaps and near-misses using technology that monitors the human factor. To explore such a technological approach to fatigue management in the naval operational environment, the current study reports feasibility and acceptability data of commercial PSM technologies in a warship environment. To our knowledge, this is the largest and most comprehensive assessment of feasibility and acceptability of commercial PSM use in an operational naval environment (Kerkamm et al., 2021; Brown et al., 2021; Conroy et al., 2022).

Despite high workload and long work hours typical of military operations, variable but generally high interest and engagement among the crews was observed. Enrollment ranged from 27 to 87% and averaged 46% of crewmembers volunteering to participate in this study. At the low end, the DDG-B crew's 27% engagement is less of an outlier and more typical of military research, even for less demanding questionnaire studies (e.g., 30% (Shattuck and Matsangas, 2017a); 25% (Matsangas and Shattuck, 2020; Newell et al., 2004)). Despite individual ships' interest in contributing to solutions for addressing crew fatigue, this underscores the operational feasibility challenges of integrating PSM technology into regular high tempo naval operations. Alternatively, the large LHD crew's 42% participation rate is a conservative estimate; study PSM supplies were limited to 300 devices per cohort, so with a nominal crew size of 1200, the actual participation rate was closer to 85% of available materials, and more consistent with the higher participation rates observed in the other crews. Importantly, for all crews, participation involved representation across the entire ship, with

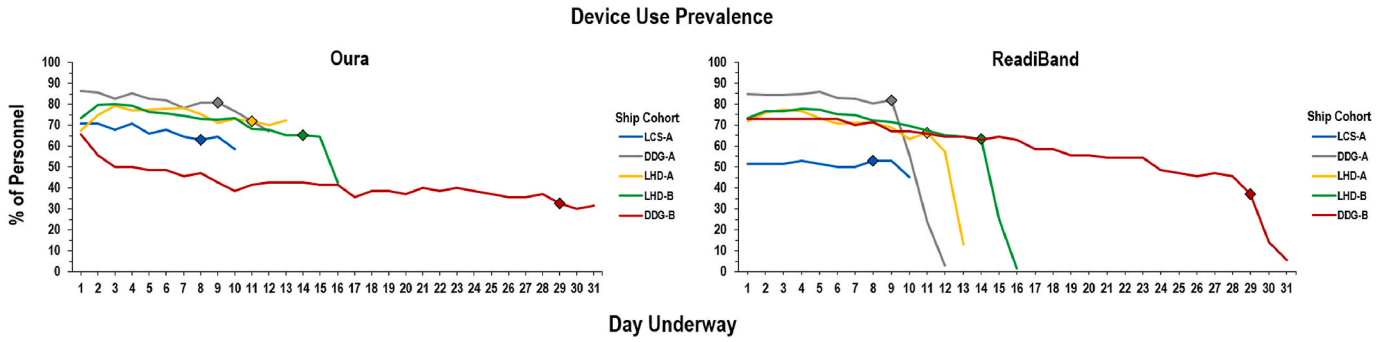


Fig. 1. Prevalence of use of Oura (left) and ReadiBand (right) devices over time at sea in five Navy ship cohorts. ◆ indicates the start day of study closeout procedures, which included collection of the ReadiBand devices.

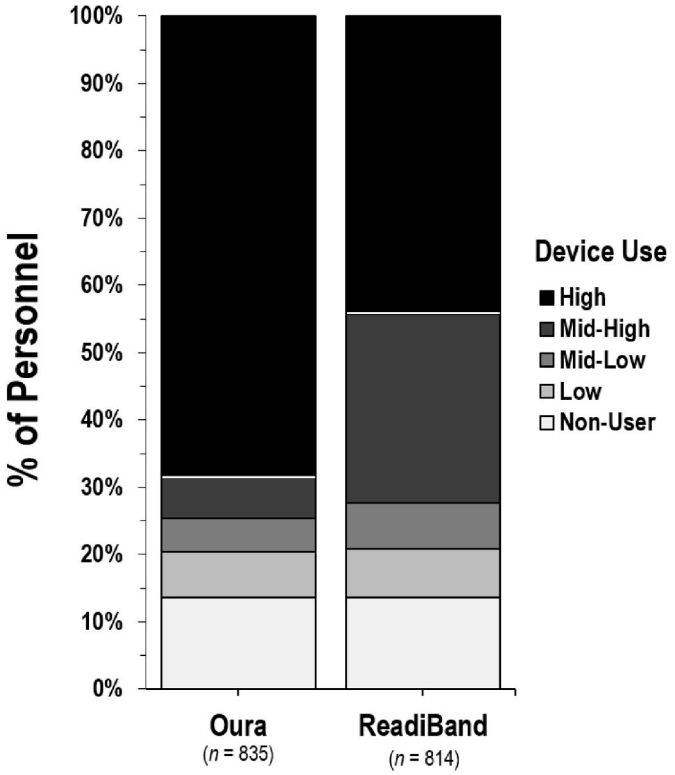


Fig. 2. Individual use rates of Oura (left) and ReadiBand (right) devices in Navy personnel while underway.

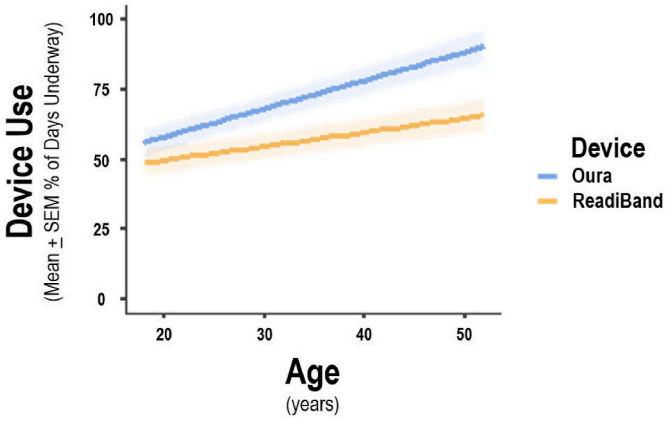


Fig. 3. Individual use rates of Oura (blue) and ReadiBand (gold) devices in Navy personnel at sea as a function of age. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

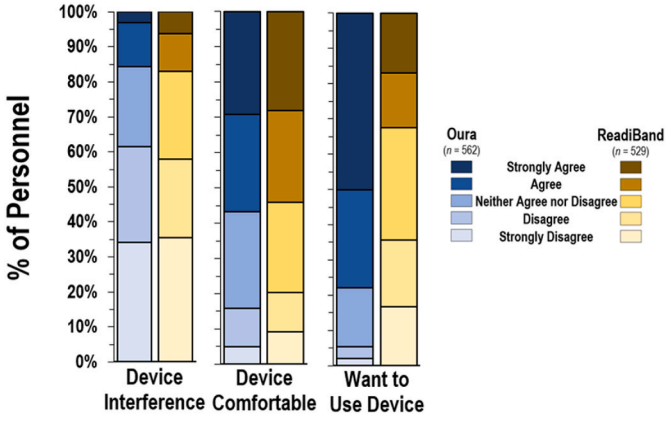


Fig. 4. Individual user experience rates for Oura (blue) and ReadiBand (gold) devices. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

no systematic exclusions based on department, paygrade, gender, age, or other demographic or operational factors. Although these findings do not cover all naval operations, they demonstrate that regular use of PSM devices for sleep and fatigue monitoring is attractive to a significant proportion of sailors, and indicates that implementing wearables is operationally feasible across a range of naval shipboard environments (Bienkowski, 1995; Claypoole et al., 2022).

Establishing technical feasibility of commercial PSM devices in operational environments is critical, but ultimately depends on the devices themselves and how they are used (Chinoy et al., 2023a). In the current investigation, technical malfunctions or failures resulting in no data capture or data transfer occurred in less than 3% of devices. This is comparable to recently observed laboratory and home use studies of these devices (OR 2.1%, RB 0.7–6.7% (Chinoy et al., 2022; Chinoy et al., 2021)). Although neither the OR nor RB devices have been validated in naval environments, laboratory and home evaluation studies have shown that both devices can accurately detect multiple sleep parameters when compared to standard research methods (e.g., PSG, ambulatory EEG, actigraphy, daily sleep logs) and sleep-tracking in other commercial devices. (Chinoy et al., 2021, 2022, 2023b). The sleep measures captured by the devices were also within ranges consistent with studies of sleep in USN personnel using self-report and research-grade actigraphy devices (Matsangas and Shattuck, 2020; Shattuck and Matsangas, 2015, 2016, 2017b), as well as known differences in sleep measure estimations between the OR and RB devices (Chinoy et al., 2022, 2023b). Detailed field validation efforts are warranted for these and any device

(s) intended for use in a warship environment, including consideration of the effects of ship motion and vibration. Taken together, these data provide initial support that commercial PSM devices can function at scale and provide interpretable sleep data for a naval fatigue management system.

Device use results offer key insights on operational acceptability. As with study participation rates, data use varied by cohort but was generally high, with the devices used in 69% of days for the OR and 71% of days with RB. In general, individual rates of OR and RB use were similar between cohorts, except for OR having much lower use than the RB in the DDG-B sample cohort. As previously mentioned, the DDG-B cohort was not able to have a research team rider on board the ship to complete check-ins and thus that cohort did not have the same pressure under the research design to wear the ring and sync the phone application daily to acquire data. This is corroborated by the fact that RB use in DDG-B tracked similarly to other cohorts and indicates that when not monitored the "set it and forget it" model device was advantageous in this cohort. While several efforts exist within the Department of Defense to explore the use of PSMs there are currently none that require PSM use as policy, and thus no opportunity to systematically assess acceptability and adoption under natural operational conditions, which would include organizational pressure. The cohort without research personnel could thus be considered a very conservative minimum estimate of PSM use, with the other ship cohorts as an intermediate pressure between no supervision and official job requirement. Although not always critical in consumer and clinical populations (Li et al., 2019; Spagnolli et al., 2014), ease of use and maintenance must be optimized for operational integration and sustained use (Howard et al., 2022) to reduce reliance on organizational pressure.

Irrespective of cohort DDG-B, individual OR device use across other cohorts was slightly greater than the RB despite having a limited battery life and the challenges of maintaining power in shipboard environments due to scarce power outlet space, which was not required of the RB. Additionally, the OR provided individual data feedback to the participants through the app interface which may have contributed to motivation to use the device despite greater effort expectancy, which can significantly contribute to technology acceptance (Venkatesh et al., 2003; Williams et al., 2015), including PSMs (Jacobs et al., 2019; Kalantari, 2017). Despite the additional effort required to sustain use of the OR versus RB, this may be an indication that devices without direct participant feedback regarding data may lead to reduced long term compliance. This is supported by the user experience data, where form factor was not a major issue, as both devices were reported to be similar in interference with daily activities and comfort, but personnel were more interested in continuing use of the OR than RB.

Demographic factors such as gender and age may influence user experience and use of PSM devices, (Williams et al., 2015; Jacobs et al., 2019; Rupp et al., 2018) although their impact on workforce and military populations is less clear (Brown et al., 2021; Wu et al., 2022). A significant relationship was found between device type and individual device use, in which use was greater for the OR than the RB. Device usage increased with age for both devices, with a stronger relationship in the OR than the RB. Although adoption of technology is typically higher in younger consumers, among USN crews, younger personnel are more likely to engage in physical work that presents a safety risk for PSMs, especially in a ring form factor. Despite high interest and engagement in these groups, frequent removal and replacement of devices adds to the use and maintenance burden, which increases both forgetting to use and ultimately abandoning the device (Attig and Franke, 2020; Jin et al., 2022). This was corroborated by exploring how responses to the user experience items varied by sociodemographic factors. A significant relationship was also found where older participants were less likely than their younger counterparts to report that their device(s) interfered with other activities. Interestingly, there was also a gender effect found in device comfort. This may be due to anatomical differences between average finger and wrist sizes between genders, although we cannot empirically confirm this with our data. Other reviews have highlighted gender effects in relation to technology user experience and attitudes, although not explicitly describing comfort (Kalantari, 2017). The impact of these demographic factors and their specificity to device type and acceptability is exploratory and should be subject to further study in the naval context (i.e., job specific duties, work center location). This may highlight that a future sleep monitoring capability that requires continuous use of PSMs will need to account for individual differences in comfort by including multiple form factor options with comparable accuracy.

4.1. Considerations and future directions

There are several factors concerning the design and approach of this research effort that should be considered. For instance, this work was completed under a voluntary research protocol with high engagement, indicating that personnel were interested in the PSMs and the research initiative to develop a sleep monitoring capability for more optimal fatigue management. However, it does not capture data from personnel who chose not to participate and their reasons for that decision, nor does it address acceptability and adoption as a job requirement. Additionally, there may be individual command climate considerations. For example, variability within leadership in the willingness and motivation to participate and encourage buy-in and support for the effort could subsequently impact wearable feasibility and acceptability by the crew. Thus, these findings cannot be generalized across the entire SURFOR fleet. In the future, topics such as privacy of data, including addressing apprehensions around leadership having access to PSM, data will need to be addressed by any organization, military or otherwise, that aims to implement PSMs in support of ORM processes (Jacobs et al., 2019; Maltseva, 2020; Menke et al., 2015; Funke et al., 2017). Individual preferences, and additional operational factors such as work duties and schedules, workload, and mission duration also warrant further exploration to better understand potential barriers to the implementation of a sleep monitoring capability onboard warships.

4.2. Conclusion

This study provides initial support that commercial PSM devices can be used to reliably collect and eventually monitor sleep data in USN warship environments. Additionally, the findings may help to inform initial feasibility estimates in similar military and civilian operational environments; however, as noted, future research must continue to assess feasibility across naval units that vary in size, mission type, and organizational structure. Several key implementation factors will need to be addressed to provide on-demand sleep monitoring from commercial PSMs. As these factors are addressed, the value of this capability to aid in the monitoring and management of fatigue can be assessed.

Declaration of interest

I am a military service member or employee of the U.S. Government. This work was prepared as part of my official duties. Title 17, U.S.C §105 provides that copyright protection under this title is not available for any work of the U.S. Government. Title 17, U.S.C §101 defines a U.S. Government work as work prepared by a military service member or employee of the U.S. Government as part of that person's official duties. This work was supported by the Military Operational Medicine Research Program under work unit no. N2010). The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, Uniformed Services University, nor the U.S. Government. Research data were derived from approved Naval Health Research Center Institutional Review Board protocol NHRC. 2021.0003.

Funding

This work was supported by the Military Operational Medicine Research Program under work unit no. N2010.

CRediT author contributions

Conceptualization: RRM, PGR, AGK, DWR. Methodology: RRM, AGK, PGR. Investigation: RRM, TBV, PGR, BJS, LRR, PHS, EDC, HNR, DAH, AGK. Data Curation: AGK, JTJ, PSG, PGR, PHS, RRM, TBV, BJS, LRR. Formal Analysis: PGR, JTJ, PSG. Visualization: PGR. Writing—Original Draft: RRM, PGR, AGK, DWR, JTJ. Writing—Review & Editing: all authors. Supervision: RRM, DWR, PGR, AGK. Project Administration: RRM, DWR. Funding Acquisition: RRM, DWR.

Data availability

The data underlying this article cannot be shared publicly because they are the property of the Department of Defense. The data may be shared on reasonable request to the corresponding author after the establishment of an institution-to-institution data sharing agreement.

Acknowledgements

We thank the leadership and crews of the participating U.S. Navy ships for their dedication and contributions. We also thank Daniel Araiza and Dr. Lily Silayeva for superb project management support, Dr. Lanny Boswell for expert consultation, and Dr. Sarah McGuire, Tara Boettcher, Joseph Cuellar, Courtney Moltzen, Dr. Alexander Posis, Catherine Watson, Sean Deering, Dr. Alice LaGoy, Josue E. Uriarte, and Andrew Thompson for their skilled technical contributions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apergo.2024.104225.

References

Attig, C., Franke, T., 2020. Abandonment of personal quantification: a review and empirical study investigating reasons for wearable activity tracking attrition. Comput. Hum. Behav. 102, 223–237. https://doi.org/10.1016/j.chb.2019.08.025, 2020/01/01/.

Barger, L.K., Runyon, M.S., Renn, M.L., et al., 2018. Effect of fatigue training on safety, fatigue, and sleep in emergency medical services personnel and other shift workers: a systematic review and meta-analysis. Prehosp. Emerg. Care 22 (Suppl. 1), 58–68.

Baumann, M.R., Gohm, C.L., Bonner, B.L., 2011. Phased training for high-reliability occupations: Live-fire exercises for civilian firefighters. Hum. Factors 53 (5), 548–557.

Bendak, S., Rashid, H.S., 2020. Fatigue in aviation: a systematic review of the literature. Int. J. Ind. Ergon. 76, 102928.

Bienkowski, M., 1995. Demonstrating the operational feasibility of new technologies: the ARPI IFDs. IEEE Expert 10 (1), 27–33. https://doi.org/10.1109/64.391960.

Bierly III, P.E., Spender, J.-C., 1995. Culture and high reliability organizations: the case of the nuclear submarine. J. Manag. 21 (4), 639–656.

Brown, J.A., Whittaker, K.S., Hawkins, S.A., Jackson Santo, T., Teyhen, D.S., 2021. Using fitness trackers to increase physical activity in a US Army health initiative. Health Educ. J. 80 (8), 883–896. https://doi.org/10.1177/00178969211026345.

Caldwell, J.A., Caldwell, J.L., Schmidt, R.M., 2008. Alertness management strategies for operational contexts. Sleep Med. Rev. 12 (4), 257–273.

Charest, J., Grandner, M.A., 2022. Sleep and athletic performance: impacts on physical performance, mental performance, injury risk and recovery, and mental health: an update. Sleep Med. Clin. 17 (2), 263–282. https://doi.org/10.1016/j.jsmc.2022.03.006, 2022/06/01/.

Chee, N.I., Ghorbani, S., Golkashani, H.A., Leong, R.L., Ong, J.L., Chee, M.W., 2021. Multi-night validation of a sleep tracking ring in adolescents compared with a research actigraph and polysomnography. Nat. Sci. Sleep 13, 177.

Chinoy, E.D., Cuellar, J.A., Huwa, K.E., et al., 2021. Performance of seven consumer sleep-tracking devices compared with polysomnography. Sleep 44 (5), zsaa291.

Chinoy, E.D., Cuellar, J.A., Jameson, J.T., Markwald, R.R., 2022. Performance of four commercial wearable sleep-tracking devices tested under unrestricted conditions at home in healthy young adults. Nat. Sci. Sleep 14, 493.

Chinoy, E.D., Markwald, R.R., 2023a. Use of technology for real-world sleep and circadian research. In: Kushida, C.A. (Ed.), Encyclopedia of Sleep and Circadian Rhythms, second ed. Academic Press, pp. 156–168.

Chinoy, E.D., Cuellar, J.A., Jameson, J.T., Markwald, R.R., 2023b. Daytime sleep-tracking performance of four commercial wearable devices during unrestricted home sleep. Nat. Sci. Sleep 15, 151–164. https://doi.org/10.2147/nss.S395732.

Claypoole, V.L., Horner, C.K., Sanchez, S.A., 2022. Augmented reality training technologies for Naval readiness: a comparison of shipboard and pier side applications. Nav. Eng. J. 134 (2), 39–47.

Connor, J., Whitlock, G., Norton, R., Jackson, R., 2001. The role of driver sleepiness in car crashes: a systematic review of epidemiological studies. Accid. Anal. Prev. 33 (1), 31–41.

Conroy, B., Silva, I., Mehraei, G., et al., 2022. Real-time infection prediction with wearable physiological monitoring and AI to aid military workforce readiness during COVID-19. Sci. Rep. 12 (1), 3797. https://doi.org/10.1038/s41598-022-07764-6. Mar 8.

de Zambotti, M., Rosas, L., Colrain, I.M., Baker, F.C., 2019. The sleep of the ring: comparison of the ŌURA sleep tracker against polysomnography. Behav. Sleep Med. 17 (2), 124–136.

Eckstein, M., 2022a. Council tests out-of-the-box ideas to boost at-sea saftey. Defense News. July 12. (Accessed 19 July 2022).

Eckstein, M., 2022b. Five years later: inside the Navy's data-driven quest to avert a future Fitzgerald or McCain collision. Defense News. June 16. (Accessed 17 June 2022).

Ferreira, J.J., Fernandes, C.I., Rammal, H.G., Veiga, P.M., 2021. Wearable technology and consumer interaction: a systematic review and research agenda. Comput. Hum. Behav. 118, 106710 https://doi.org/10.1016/j.chb.2021.106710, 2021/05/01/.

Funke, G.J., Best, C., Menke, L.E., Strang, A.J., 2017. Warfighter acceptance of future physio-behavioral monitoring and augmentation: update. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 61 (1), 1151–1155. https://doi.org/10.1177/1541931213601772.

Gottschall, S., Guérin, E., 2021. Organizational and non-organizational risk and resilience factors associated with mental health and well-being in the Royal Canadian Navy. Curr. Psychol. 1–15.

Gould, K.S., Røed, B.K., Koefoed, V.F., Bridger, R.S., Moen, B.E., 2006. Performance-shaping factors associated with navigation accidents in the Royal Norwegian Navy. Mil. Psychol. 18 (Suppl. 1), S111–S129.

Grandou, C., Wallace, L., Fullagar, H.H.K., Duffield, R., Burley, S., 2019. The effects of sleep loss on military physical performance. Sports Med. 49 (8), 1159–1172. https://doi.org/10.1007/s40279-019-01123-8, 2019/08/01.

Howard, J., Murashov, V., Cauda, E., Snawder, J., 2022. Advanced sensor technologies and the future of work. Am. J. Ind. Med. 65 (1), 3–11. https://doi.org/10.1002/ajim.23300. Jan.

Islam, R., Anantharaman, M., Khan, F., Garaniya, V., 2020. A review of human error in marine engine maintenance. TransNav: Int. J. Mar. Navigation and Safety of Sea Transport. 14 (1).

Jacobs, J.V., Hettinger, L.J., Huang, Y.H., et al., 2019. Employee acceptance of wearable technology in the workplace. Appl. Ergon. 78, 148–156. https://doi.org/10.1016/j.apergo.2019.03.003. Jul.

Jameson, J.T., Markwald, R.R., Kubala, A.G., et al., 2022. Sleep deficiency, operational fatigue and the interplay of compromising factors: analysis to aid in fatigue management. J. Sleep Res., e13788 https://doi.org/10.1111/jsr.13788. Nov 27.

Jin, D., Halvari, H., Maehle, N., Olafsen, A.H., 2022. Self-tracking behaviour in physical activity: a systematic review of drivers and outcomes of fitness tracking. Behav. Inf. Technol. 41 (2), 242–261. https://doi.org/10.1080/0144929X.2020.1801840, 2022/01/25.

Kalantari, M., 2017. Consumers' adoption of wearable technologies: literature review, synthesis, and future research agenda. Int. J. Technol. Market. 12, 274–307. https://doi.org/10.1504/IJTMKT.2017.10008634, 01/01.

Kancherla, B.S., Upender, R., Collen, J.F., et al., 2020. Sleep, fatigue and burnout among physicians: an American Academy of Sleep Medicine position statement. J. Clin. Sleep Med. 16 (5), 803–805.

Kerkamm, F., Dengler, D., Eichler, M., et al., 2021. Measurement methods of fatigue, sleepiness, and sleep behaviour aboard ships: a systematic review. Int. J. Environ. Res. Publ. Health (1), 19. https://doi.org/10.3390/ijerph19010120. Dec 23.

Lewis, J.R., 2018. The system usability scale: past, present, and future. Int. J. Hum. Comput. Interact. 34 (7), 577–590. https://doi.org/10.1080/10447318.2018.1455307, 2018/07/03.

Li, J., Ma, Q., Chan, A.H., Man, S.S., 2019. Health monitoring through wearable technologies for older adults: smart wearables acceptance model. Appl. Ergon. 75, 162–169. https://doi.org/10.1016/j.apergo.2018.10.006. Feb.

Maltseva, K., 2020. Wearables in the workplace: the brave new world of employee engagement. Bus. Horiz. 63 (4), 493–505. https://doi.org/10.1016/j.bushor.2020.03.007, 2020/07/01/.

Manuele, F.A., 2013. Preventing serious injuries & fatalities: time for a sociotechnical model for an operational risk management system. Prof. Saf. 58 (5), 51–59.

Matsangas, P., Shattuck, N.L., 2020. Sleep quality, occupational factors, and psychomotor vigilance performance in the U.S. Navy sailors. Sleep 43 (12). https://doi.org/10.1093/sleep/zsaa118.

Matsangas, P., Shattuck, N.L., 2021. Habitability in berthing compartments and well-being of sailors working on US Navy surface ships. Hum. Factors 63 (3), 462–473.

Menke, L., Best, C., Funke, G.J., Strang, A.J., 2015. Warfighter acceptance of future physiological monitoring and augmentation. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 59, 125–129.

National Transportation Safety Board. https://www.ntsb.gov/safety/safety-recs/reclletters/M90_26_31A.pdf.

Newell, C.E., Rosenfeld, P., Harris, R.N., Hindelang, R.L., 2004. Reasons for nonresponse on U.S. Navy questionnaires: a closer look. Mil. Psychol. 16, 265–276. https://doi.org/10.1207/s15327876mp1604_4.

Peets, A., Ayas, N.T., 2012. Restricting resident work hours: the good, the bad, and the ugly. Crit. Care Med. 40 (3), 960–966. https://doi.org/10.1097/CCM.0b013e3182413bc5.

Reifman, J., Priezjev, N.V., Vital-Lopez, F.G., 2023. Can we rely on wearable sleep-tracker devices for fatigue management? Nov 8 Sleep zsad288. https://doi.org/10.1093/sleep/zsad288. Epub ahead of print.

Rentz, L.E., Ulman, H.K., Galster, S.M., 2021. Deconstructing commercial wearable technology: contributions toward accurate and free-living monitoring of sleep. Sensors 21 (15), 5071.

Roberts, K.H., 1990. Some characteristics of one type of high reliability organization. Organ. Sci. 1 (2), 160–176.

Rosekind, M.R., Gander, P.H., Miller, D.L., et al., 1994. Fatigue in operational settings: examples from the aviation environment. Hum. Factors 36 (2), 327–338.

Rupp, M.A., Michaelis, J.R., McConnell, D.S., Smither, J.A., 2018. The role of individual differences on perceptions of wearable fitness device trust, usability, and motivational impact. Appl. Ergon. 70, 77–87. https://doi.org/10.1016/j.apergo.2018.02.005. Jul.

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. J. Sleep Res. 30 (6), e13397.

Sánchez-Beaskoetxea, J., Basterretxea-Iribar, I., Sotés, I., Machado, MdlMM., 2021. Human error in marine accidents: is the crew normally to blame? Mar. Transport Res. 2, 100016 https://doi.org/10.1016/j.martra.2021.100016, 2021/01/01/.

Shattuck, N.L., Matsangas, P., 2015. A 6-month assessment of sleep during naval deployment: a case study of a commanding officer. Aerosp Med. Hum. Perform 86 (5), 481–485. https://doi.org/10.3357/amhp.4140.2015. May.

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. https://doi.org/10.1080/00140139.2015.1073794, 2016/05/03.

Shattuck, N.L., Matsangas, P., 2017a. Sunlight exposure, work hours, caffeine consumption, and sleep duration in the naval environment. Aerospace Med. Hum. Performance 88 (6), 579–585.

Shattuck, N.L., Matsangas, P., 2017b. Sunlight exposure, work hours, caffeine consumption, and sleep duration in the naval environment. Aerosp Med. Hum. Perform 88 (6), 579–585. https://doi.org/10.3357/amhp.4721.2017. Jun 1.

Spagnolli, A., Guardigli, E., Orso, V., Varotto, A., Gamberini, L., 2014. Measuring User Acceptance of Wearable Symbiotic Devices: Validation Study across Application Scenarios.

Stern, H.S., Blower, D., Cohen, M.L., et al., 2019. Data and methods for studying commercial motor vehicle driver fatigue, highway safety and long-term driver health. Accid. Anal. Prev. 126, 37–42.

Stone, J.D., Rentz, L.E., Forsey, J., et al., 2020. Evaluations of commercial sleep technologies for objective monitoring during routine sleeping conditions. Nat. Sci. Sleep 12, 821.

Taylor, A.H., Dorn, L., 2006. Stress, Fatigue, Health and Risk of Road Traffic Accidents Among Professional Drivers. The Contribution of Physical Inactivity.Annu. Rev. Public Health. 7, 371–391. https://doi.org/10.1146/annurev.publhealth.27.021405.102117.

U.S. Department of Defense. *2022 Demographics Profile of the Military Community*. https://www.militaryonesource.mil/data-research-and-statistics/military-community-demographics/2022-demographics-profile/.

U.S. Government Accountability Office, 2021. *Navy Readiness: Additional Efforts Are Needed to Manage Fatigue, Reduce Crewing Shortfalls, and Implement Training*. Report GAO-21-366. DC: Government Accountability Office, Washington.

Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D., 2003. User acceptance of information technology: toward a unified view. MIS Q. 27, 425–478. https://doi.org/10.2307/30036540.

Williams, M., Rana, N., Dwivedi, Y., 2015. The unified theory of acceptance and use of technology (UTAUT): a literature review. J. Enterprise Inf. Manag. 28, 443–488. https://doi.org/10.1108/JEIM-09-2014-0088, 04/13.

Williamson, A., Lombardi, D.A., Folkard, S., Stutts, J., Courtney, T.K., Connor, J.L., 2011. The link between fatigue and safety. Accid. Anal. Prev. 43 (2), 498–515. https://doi.org/10.1016/j.aap.2009.11.011, 2011/03/01/.

Wu, R., Yiu, T., Jelodar, M., 2022. Real-time employee monitoring technologies in the construction sector – effect, readiness and theoretical perspectives: the case of New Zealand. IOP Conf. Ser. Earth Environ. Sci. 1101, 082010 https://doi.org/10.1088/1755-1315/1101/8/082010, 11/01.