Assessment of a Novel Watchstanding Schedule on an Operational us Navy Vessel

Nita Lewis Shattuck, Panagiotis Matsangas and Lauren Waggoner
Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2014 58: 2265
DOI: 10.1177/1541931214581472

ASSESSMENT OF A NOVEL WATCHSTANDING SCHEDULE ON AN OPERATIONAL US NAVY VESSEL

Nita Lewis Shattuck, Panagiotis Matsangas and Lauren Waggoner
Operations Research Department, Naval Postgraduate School, Monterey, CA

One watchstanding schedule used in the US Navy is the modified 6-hr on/18-hr off, a 6-hr watch followed by an 18-hr non-watch period. In this four-section watchbill, rather than standing the same 6-hour period each day, two 3-hr watches are "dogged", resulting in a day shorter than 24-hrs in length. We assessed the 6-hr on/18-hr off schedule to determine its affordance of rest/sleep and psychomotor vigilance, comparing it to a 3-hr on/9-hr off schedule. Results show the 3-hr on/9-hr off schedule is superior, affording 0.63 more hours of rest daily ($p=0.054$), 0.89 more hours of sleep daily ($p=0.024$), and decreased variability in psychomotor vigilance (i.e., reaction time and lapses) compared to the standard schedule ($p<0.05$). Variability of daily rest and sleep is also reduced ($p<0.05$). Subjective assessments between the 3-hr on/9-hr off and seven other watch schedules showed that participants preferred the 3-hr on/9-hr off schedule. The 3-hr on/9-hr off schedule yields better sleep hygiene, more stable performance and is well-accepted by crewmembers.

INTRODUCTION

Research has established that members of the military get inadequate sleep (Miller, Matsangas, & Kenney, 2012). Sleep deprivation in the military is traced back to the ancient Greeks (Shay, 1998). While all branches of the military seem steeped in this long tradition of sacrificing sleep, the "24/7" nature of continuous operations in the current defense climate has further exacerbated the problem. The naval environment, in particular, is characterized by sleep problems, sleep deprivation and increased levels of fatigue (Howarth, Pratt, & Tepas, 1999).

Historically, US Navy (USN) ships operate continually with qualified personnel standing watch at various critical locations on the ship – typically, the bridge of the ship, the combat center, and the engineering spaces. This manning requirement brings with it the need to have a crew sized such that crewmembers have time off watch for other work and rest activities. The watchstanding schedule depends on the organizational culture, prior experiences of the command leadership and the number of qualified sailors available to stand watch, a major limitation on ships with small crews.

The USN often uses a rotating watchbill of 5 hours in length with 10 hours off watch; it is the responsibility of the officer of the watch to ensure that watch standers are rotated frequently enough to stand an effective watch (Department of the Navy, 2012). Given the availability of personnel, the watch itself, and other daily activities, a number of fixed and rotating watch systems are used, e.g., 6-hr on/6-hr off, 12-hr on/12-hr off, 4-hr on/8-hr off, 6-hr on/18-hr off, or 3-hr on/9-hr off. Schedules like the 5-hr on/10-hr off and the 5-hr on/15-hr off result in circadian misalignment equating to a 15 or 20-hour day without weekends or time for recovery. In contrast to the modified 6-hr on/18-hr off watch schedule, we have proposed the adoption of a 3-hr-on/9-hr off circadian-aligned schedule based on a 4-section watchbill where sailors stand 3-hour watches that commence every 12 hours.

Watchbills are characterized by the number of watch sections needed to continually stand watch and are typically described as 2, 3, or 4 section watchbills, requiring 2, 3 or 4 individuals, respectively, to stand watch at a particular location to cover the entire day. Efforts have been made to reduce manning in an attempt to align costs to a shrinking defense budget. Studies on naval vessels showed that watch schedules traditionally used at sea lead to sleep deprivation, sleep fragmentation, sub-optimal performance and worrisome levels of cognitive effectiveness (Paul, Ebisuzaki, McHarg, Hursh, & Miller, 2012; Rutenfranz et al., 1988).

The traditional 6-hr on/18-hr off watch schedule has 4 sections, in which a 6-hr watch period is followed by an 18-hr non-watch period. Watchstanders stand a "fixed" watch, reporting for watch at the same time each day. In the modified 6-hr on/18-hr off schedule, there are two 3-hr watches ("dog watches") from 00:00 to 03:00 and from 03:00 to 06:00. These "dog" watches cause the schedule to rotate backwards, with watchstanders reporting earlier on each successive day. An serious consequence of the modified 6/18 is that each day is shorter than a normal 24-hour day. The modified 6-hr on/18-hr off schedule is shown in Figure 1.

Day	00:00 - 03:00	03:00 - 06:00	06:00 - 12:00	12:00 - 18:00	18:00 - 24:00
Day 1	Section 1	Section 2	Section 3	Section 4	Section 1
Day 2	Section 2	Section 3	Section 4	Section 1	Section 2
Day 3	Section 3	Section 4	Section 1	Section 2	Section 3
Day 4	Section 4	Section 1	Section 2	Section 3	Section 4
Day 5	Section 1	Section 2	Section 3	Section 4	Section 1
Day 6	Section 2	Section 3	Section 4	Section 1	Section 2
Day 7	Section 3	Section 4	Section 1	Section 2	Section 3
Day 8	Section 4	Section 1	Section 2	Section 3	Section 4
Day 9	Section 1	Section 2	Section 3	Section 4	Section 1
Day 10	Section 2	Section 3	Section 4	Section 1	Section 2
Day 11	Section 3	Section 4	Section 1	Section 2	Section 3

Figure 1. Typical 4-section modified 6-hr on/18-hr off watch schedule.

Not subject to U.S. copyright restrictions. DOI 10.1177/1541931214581472

The objective of this study was to assess the sleep patterns and psychomotor vigilance of crewmembers while using the modified 6-hr on/18-hr off watchstanding schedule. We sought to determine how the performance in the 6/18 compares to that of sailors working a 3-hr on/9-hr off schedule.

METHODS

Participants

Sixty crewmembers of the Arleigh Burke Flight I (approximately 9000 tons) class destroyer volunteered to participate in the underway sleep and performance assessment. Of those, 34 individuals (26 males, 8 females, 11 officers and 23 enlisted crewmembers) were included in this analysis. To be included in the study, participants were required to be watchstanders on the modified 6-hr on/18-hr off watch schedule, and to have valid actigraphic sleep recordings and psychomotor vigilance task data. Participants were on average 26.9±6.7 years old, and had on average 6.31±6 years of military service.

Equipment and Instruments

The Motionlogger Watch (Ambulatory Monitoring, Inc.-AMI; Ardsley, NY) was used to collect actigraphic recordings of sleep in the study. Actigraphic data were collected in one-minute epochs. AMI data (collected in the Zero-Crossing Mode) were scored using Action W version 2.7.2155 software. The Cole-Kripke algorithm with rescoring rules was used in the analysis. For sleep scoring, all values were set to the default values. In addition to wearing a wrist activity monitor, all participants were asked to complete a daily activity log, documenting their daily routine. The logs covered each 24-hour period in 30-minute intervals.

The Epworth Sleepiness Scale (ESS) was used to assess average daytime sleepiness (Johns, 1991). In filling out the ESS, an individual uses a 4-item Likert scale to rate the likelihood of dozing off or falling asleep in eight different everyday situations. Scoring of the answers ranges from 0 to 3, with 0 being "would never doze," 1 is "slight chance of dozing," 2 is "moderate chance of dozing," and 3 denotes "high chance of dozing." The respondent is instructed to "rate according to his/her usual way of life in recent times". Responses are summed to arrive at the total ESS score. A score of more than 10 reflects above normal daytime sleepiness and suggests the need for further evaluation (Johns, 1992). The Morningness-Eveningness (M-E) Scale (Horne & Östberg, 1976), was also administered to assess participants' circadian chronotype, an attribute of humans indicating their preference for waking earlier or later in the day. The M-E scale includes 19 multiple-choice questions. Scores range from 16 to 86, with scores less than 42 corresponding to evening-preference chronotypes and scores higher than 58 indicating morning-preference chronotypes.

The evaluation of the participants' subjective sleep quality was assessed with the Pittsburgh Sleep Quality Index (PSQI) scores (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989). The PSQI includes 18 questions that yield seven component scores (sleep quality, sleep latency, duration, sleep efficiency, sleep disturbances, sleep medication use, and daytime dysfunction) rated from 0 (better) to 3 (worse). The total score, ranging from 0 (better) to 21 (worse), is the summation of the component scores. Individuals with a PSQI total score $\leq$5 are characterized as good sleepers, whereas scores >5 are associated with poor sleep quality.

Performance data were collected with the Psychomotor Vigilance Task (PVT) (Dinges & Powell, 1985). PVT performance is not only affected by sleep loss, it has also been shown that the PVT is sensitive to circadian rhythmicity (Dinges et al., 1997; Doran, Van Dongen, & Dinges, 2001; Durmer & Dinges, 2005; Jewett, Dijk, Kronauer, & Dinges, 1999; Wyatt et al., 1997). The PVT involves a simple reaction time test where participants are required to press a response button as soon as the stimulus appears on the screen. Because of its simplicity, PVT has very minor learning effects, which are reached in 1 to 3 trials (Dinges et al., 1997; Jewett et al., 1999; Kribbs & Dinges, 1994; Rosekind et al., 1994). The nominal inter-stimulus interval (ISI), defined as the period between the last response and the appearance of the next stimulus, randomly ranges from 2 to 10 sec. The standard version of the PVT has a duration of 10 minutes (Loh, Lamond, Dorrian, Roach, & Dawson, 2004). However, shortened version have also shown their utility to assess sleep deprivation effects (Basner & Dinges, 2011; Loh et al., 2004). Operational demands prevented the use of the 10-minute version in this study; consequently, we used a 3-minute version of PVT administered on the AMI actigraphs. ISI ranged from 2 to 10 seconds. A red backlight appeared for one second and the letters "PUSH" were used as visual stimuli; the response time was then displayed in milliseconds.

Procedures

The quasi-experimental study was conducted onboard a USN Arleigh Burke class destroyer during a transit from Hawaii to San Diego between May 24 and May 30, 2013. The study protocol was approved by the NPS Institutional Review Board and informed consent was obtained from all participants at the beginning of the study.

Data cleaning and reduction

The primary source of sleep data was from the actigraphic recordings from the individual participants. However, since there were occasional gaps and discrepancies in the actigraphic recordings, we also evaluated the self-reported activity logs, comparing the actigraphy data with these logs. The sleep logs also assisted in the determination of start and end time of the sleep intervals. Based on this comparison, we adjusted the start and end of sleep episodes in the actigraphy data.

Psychomotor vigilance task (PVT) performance was collected using a version of the PVT which was conveniently integrated into the AMI Motionloggers. The duration of each PVT trial was 3 minutes with minimum interstimulus interval (ISIMin) of 2 seconds, and maximum interstimulus interval (ISIMax) of 10 seconds. PVT data were analyzed based on the metrics of Basner and Dinges (2011) for individuals with chronic sleep deprivation. Specifically, a PVT response was regarded as valid if the RT was ≥ 100 ms. Responses without a stimulus or RTs < 100 ms were identified as false starts. Two levels of lapses were defined as RTs ≥ 350 and RTs ≥ 500 ms.

Analytical strategy

Statistical analysis was conducted with a statistical software package (JMP Pro 9; SAS Institute; Cary, NC). To assess the modified 6-hr on/18-hr off schedule, we used the daily rest and sleep amount, PVT performance (the mean reaction time (RT), the median RT, the mean response speed (1/RT), fastest 10% RT (i.e., 10^{th} percentile of RT), slowest 10% of 1/RT (i.e., 10^{th} percentile of 1/RT), the percentage of lapses, the percentage of false starts, and the percentage of the sum of lapses and false starts. Data are presented as mean (M) ± standard deviation (SD) or median (MD) as appropriate. Significance level was set at $p < 0.05$. Wilcoxon Rank Sum test was used for comparisons. Correlation analysis was performed using the non-parametric Spearman's rho. Variability between groups was assessed using Levene's test.

Data were grouped by participant. Specifically for PVT metrics, we averaged responses by participant, not by trial. The reason for this decision is the number of responses in each PVT 3-minute trial, which is approximately 20. Calculating means, percentiles and lapses/false starts by trial would introduce one more level or error, when aggregating from the response to the trial level, and then to the participant level. Furthermore, the focus of this study was to evaluate the modified 6-hr on/18-hr off watch schedule over the entire data collection period. Aggregating by participant over the entire data collection period provided a better reflection of the overall effect of the watch schedule on the participants.

Initially, a descriptive analysis was performed focusing on sleep intervals and daily rest and sleep amounts, followed by an analysis by participant. Then, we assessed the differences between the modified 6-hr on/18-hr off and the 3-hr on/9-hr off watch schedules in terms of daily rest, daily sleep, and PVT performance metrics. The unpublished data for the 3-hr on/9-hr off comparison are based on a data collection done in December of 2012 on the crew of a similar ship.

RESULTS

The average morningness – eveningness preference (MEP) score was 54±6.5. Two participants were moderately morning type, two were moderately evening type, while 27 were neither morning nor evening type. The average PSQI Global score was 9.14±2.73 ranging from 5 to 15. PSQI scores suggested that none of the participants in the modified 6-hr on/18-hr off watch schedule were "good sleepers" (PSQI score < 5).

The average Epworth Sleepiness Scale (ESS) score for the participants was 10.4 ± 4.01 (MD = 10) ranging from 4 to 19. ESS scores suggested that 16 participants (47.1%) showed elevated daytime sleepiness (ESS score > 10) (Johns, 1991).

This analysis is based on a total of 372 sleep episodes from 34 participants. On average, each participant obtained 10.9 ± 3.05 sleep episodes (MD = 10), ranging from 6 to 20. The average fragmentation index was 1.86 ± 0.518 (MD=1.70) ranging from 1.02 to 3.40. The 372 sleep episodes correspond to 204 sleep-days with each participant providing 6 days of sleep data. Table 1 shows the daily rest and sleep amounts averaged by participant.

Table 1. Daily rest and sleep amount by participant in hours

	M	SD	MD	Min	Max
Rest	6.62	1.66	6.67	2.58	10.5
Sleep	5.65	1.63	5.73	1.83	9.52

It is notable that all but one of the crewmembers received on average less than 8 hours of sleep daily, while 59% of crewmembers received less than 6 hours. On average, participants accrued a sleep debt of approximately 2.35 hours each day. In the 6-day study period, crewmembers accumulated approximately 14 hours of chronic sleep deficit as calculated using the recommended 8 hours of sleep each day for healthy adults.

Table 2 describes the ten PVT performance metrics used in this study.

Table 2. PVT metrics

PVT metric	M	Min	Max
Mean RT, [ms]	349±140	203	730
Median RT, [ms]	269±76	172	500
Mean 1/RT	3.91±0.86	1.93	5.81
Fastest 10% RT, [ms]	193±48.9	133	356
Slowest 10% 1/RT	2.37±0.81	0.552	4.29
False Starts (FS), [%]	2.25±3.23	0%	18.9%
Lapses 500ms, [%]	10.4±12.4	1.02%	49.9%
Lapses 355ms, [%]	21.2±20.6	1.53%	90.4%
Lapses 500ms+FS, [%]	12.6±13.1	2.04%	50.5%
Lapses 355ms+FS, [%]	23.5±20.8	2.55%	90.4%

Lastly, we assessed the differences between the modified 6-hr on/18-hr off (n = 34) and the 3-hr on/9-hr off (n = 24) watch schedules in terms of daily rest, daily sleep, and PVT performance metrics (mean RT, lapses and false alarms). The modified 6-hr on/18-hr off afforded fewer rest and sleep opportunities than the 3-hr on/9-hr off. Although the differences in the analyzed PVT metrics were not statistically significant, crewmembers in the modified 6-hr on/18-hr off have significantly greater variability in psychomotor performance in terms of reaction times and lapses combined with false starts compared to their peers in the 3-hr on/9-hr off. Table 3 describes these results.

Table 3. Comparison of rest, sleep, and PVT performance metrics between the modified 6-hr on/18-hr off and the 3-hr on/9-hr off watchstanding schedules

PVT metric	3/9 M ± SD	6/18mod M ± SD	Level Δ [A] p-value	Variability Δ [B] p-value
Daily rest, hrs	7.25±0.8	6.62±1.7	0.054	0.013
Daily sleep, hrs	6.54±0.8	5.65±1.6	0.024	0.010
Mean RT, [ms]	365±105	349±140	-	-
Median RT, [ms]	253±35	269±76	-	0.033
Lapses 500ms+FS, [%]	9.23±4.7	12.6±13.1	-	0.010
Lapses 355ms+FS, [%]	18.7±9.5	23.5±20.8	-	0.022

Inclusion criterion: $p<0.10$; [A] Comparisons using Wilcoxon Rank Sum test; [B] Variability differences using Levene's test

These results are further depicted in the following figures.

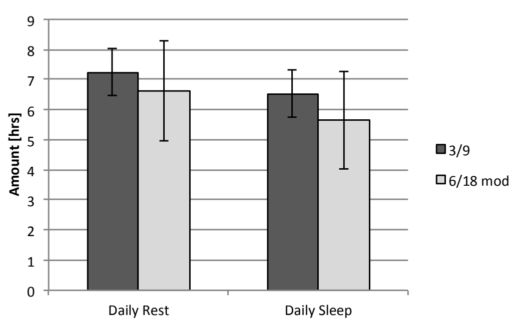

Figure 2. Daily rest and sleep amounts by watch schedule. Vertical bars represent one standard deviation.

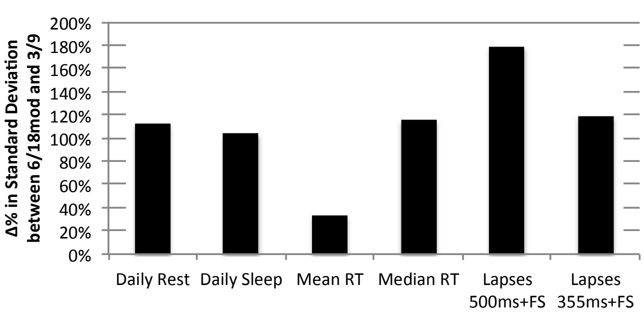

Figure 3. Percentage-wise differences in standard deviation between the modified 6/18 and the 3-hr on/9-hr off watch schedules.

After the completion of the study, participants filled out a questionnaire that compared seven watch schedules (5-hr on/10-hr off, 5-hr on/15-hr off, 6-hr on/6-hr off, 12-hr on/12-hr off, 6-hr on/12-hr off, 3-hr on/9-hr off, 4-hr on/8-hr off) with their modified 6-hr on/18-hr off. They were asked to respond to the question "Compared to my current schedule, the [watch schedule] is …" using a 3 point Likert scale (Worse "1", Same as "2", Better "3"). Results show that the participants evaluate the 6-hr on/6-hr off as the worst schedule and the 3-hr on/9-hr off as the best. Figure 4 shows the integrated results by watchstanding schedule.

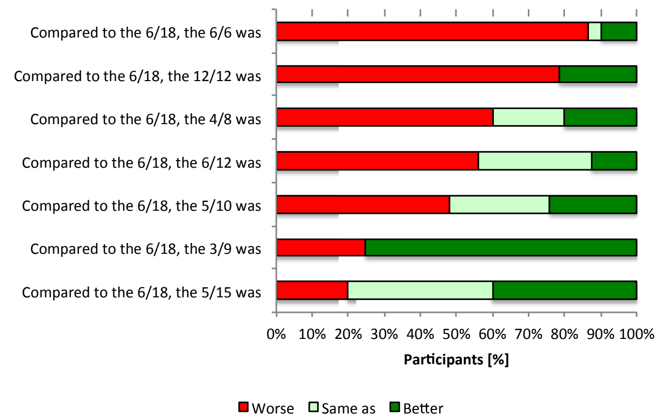

Figure 4. Subjective evaluations of watch schedules.

DISCUSSION

Results indicate that crewmembers on the 3-hr on/9-hr off watchstanding schedule received 0.63 hours more daily rest compared to crewmembers on the modified 6-hr on/18-hr off schedule ($p = 0.054$). Additionally, sailors working on the 3-hr on/9-hr off received 0.89 hours more daily sleep than those on the modified 6-hr on/18-hr off schedule ($p = 0.024$). Although average PVT performance in terms of reaction time and lapses did not differ significantly between the two schedules, the 6-hr on/18-hr off watchstanding schedule had significantly larger variability in crewmembers' performance compared to 3-hr on/9-hr off schedule ($\Delta(SD) > 100\%$, $p < 0.05$). The same phenomenon was observed in daily rest and sleep amount ($\Delta(SD) > 100\%$, $p < 0.05$). Furthermore, subjective assessments among the modified 6-hr on/18-hr off and seven other watch schedules (5/10, 5/15, 6/6, 12/12, 6/12, 3-hr on/9-hr off, 4/8) showed that participants subjectively rated the 3-hr on/9-hr off schedule as the highest of the seven schedules.

In summary, the 3-hr on/9-hr off schedule yielded better sleep hygiene, more stable performance and was well-accepted by crewmembers. The surface navy community should consider re-evaluating its watchstanding practices. This study suggests that watchstanding schedules based on sound human performance and ergonomics principles may lead to better performance in the operational environment that are better accepted by crewmembers.

REFERENCES

Basner, M., & Dinges, D. F. (2011). Maximizing sensitivity of the Psychomotor Vigilance Test (PVT) to sleep loss. *Sleep, 34*(5), 581-591.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a New Instrument for Psychiatric Practice and Research. *Journal of Psychiatric Research, 28*(2), 193-213.

Department of the Navy. (2012). Standard organization and regulations of the U.S. Navy - OPNAV Instruction 3120.32D (pp. 452). Washington, DC: Department of the Navy.

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., . . . Pack, A. I. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep, 20*(4), 267-277.

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers, 17*(6), 652-655.

Doran, S. M., Van Dongen, H. P. A., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: Evidence of state instability. *Archives Italiennes de Biologie, 139*(3), 253-267.

Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive Consequences of Sleep Deprivation. *Seminars in Neurology, 25*(1), 117-129.

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

Howarth, H. D., Pratt, J., & Tepas, D. I. (1999). Do maritime crew members have sleep disturbances? *International Journal of Occupational and Environmental Health, 5*(2), 95-100.

Jewett, M. E., Dijk, D. J., Kronauer, R. E., & Dinges, D. F. (1999). Dose-response relationship between sleep duration and human psychomotor vigilance and subjective alertness. *Sleep, 22*(2), 171-179.

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

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

Kribbs, N. B., & Dinges, D. F. (1994). Vigilance decrement and sleepiness. In J. R. Harsh & R. D. Ogilvie (Eds.), *Sleep Onset Mechanisms* (pp. 113-125). Washington, DC: American Psychological Association.

Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers, 36*(2), 339-346.

Miller, N. L., Matsangas, P., & Kenney, A. (2012). The Role of Sleep in the Military: Implications for Training and Operational Effectiveness. In J. H. Laurence & M. D. Matthews (Eds.), *The Oxford Handbook of Military Psychology* (pp. 262–281): Oxford University Press.

Paul, M. A., Ebisuzaki, D., McHarg, J., Hursh, S. R., & Miller, J. C. (2012). An assessment of some watch schedule variants used in Canadian Patrol Frigates (pp. 105). Toronto, Canada: Defence Research and Development.

Rosekind, M. R., Graeber, R. C., Dinges, D. F., Connel, L. J., Rountree, M. S., & Spinweber, C. L. (1994). Crew factors in flight operations IX: effects of planned cockpit rest on crew performance and alertness in long-haul operations. Moffett Field, CA: NASA Ames Research Center.

Rutenfranz, J., Plett, R., Knauth, P., Condon, R., De Vol, D., Flethcher, N., . . . Colquhoun, W. P. (1988). Work at sea: a study of sleep, and circadian rhythms in physiological and psychological functions, in watchkeepers on merchant vessels. *International Archives of Occupational and Enivornmental Health, 60*(5), 331-339.

Shay, J. (1998). Ethical standing for commander self-care: The need for sleep. *Parameters, 28*(2), 93-105.

Wyatt, J. K., Dijk, D. J., Ronda, J. M., Jewett, M. E., Powell, J. W., Dinges, D. F., & Czeisler, C. A. (1997). Interaction of circadian- and sleep/wake homeostatic-processes modulate psychomotor vigilance test (PVT) performance. *Sleep Research, 26*, 759.