Members of the military get inadequate sleep due to a variety of reasons. Reduced manning, extended work hours, shiftwork schedules that result in circadian misalignment -- all of these factors contribute to the sleep debt and degraded alertness observed in much of the military population. The issue of watchstanding schedules, performance, and alertness is of critical importance to the US military and is the focus of the current study. Based on a sample of active duty military members (N=75), this study had two goals. First, to conduct a field-based monitoring of the sleep and performance of military personnel while performing their duties. Second, to create and validate optimal recommendations based on the results of this empirical study. Participants wore actigraphs over a two-week period, completed daily activity logs, and took three-minute reaction time tests before and after standing watch on their regular schedules. Participants worked on a 2-day on/2-day off schedule, either in 3-section 8-hour shifts, or 2-section 12-hour shifts. Although there were no significant differences in the sleep amounts between the two schedules, results showed that participants on 8-hr shifts had fewer errors and less variable reaction time performance than those working 12-hr shifts. The 8-hr group reported better sleep quality, too. Our results suggest that the 8-hour schedule is better than the 12-hour schedule in terms of sleep and performance but may be more difficult to be applied. This study clearly shows the difficulty of implementing a specific watchstanding schedule in operational environments overloaded with unplanned, and irregular operational duties.
Equipment and Instruments

The pre-study questionnaire included demographic information and five standardized questionnaires. The Epworth Sleepiness Scale (ESS) was used to assess average daytime sleepiness (Johns, 1991). The Insomnia Severity Index (ISI) was used to assess the severity of both nighttime and daytime components of insomnia (Bastien, Vallieres, & Morin, 2001). Participants’ sleep history was assessed using the Pittsburg Sleep Quality Index (PSQI) (Buysse, Reynolds III, Monk, Berman, & Kupfer, 1989). To measure mood states and assess changes in mood, participants filled out the Profile of Mood State (POMS) scale (McNair, Lorr, & Droppelman, 1971). The questionnaire assesses various dimensions of mood using six subscales: anger - hostility, confusion - bewilderment, depression, fatigue, tension – anxiety, and vigor – activity. The Total Mood Disturbance (TMD) score was derived by adding five of the subscales and subtracting the score for vigor.

The end-of-study questionnaire included the ESS, the ISI, and the POMS. The end-of-study questionnaire also included three open-ended questions (“What did you like most about your current watch schedule,”“What did you like least about your current watch schedule,” “What advice would you give to others who would like to improve their watchstanding schedules.”)

Supplemented with activity logs, sleep data were collected with the Ambulatory Monitoring, Inc. (AMI) Motionlogger Watch in 1-minute epochs. AMI data (collected in the Zero-Crossing Mode) were scored using Action W version 2.7.2155 software. The Cole-Kripke algorithm with rescoring rules was used to score sleep; the criterion for sleep and wake episode length was set at five minutes. The sleep latency criterion was no more than 1-minute wake in 20 minutes period.

Performance data were collected using a 3-minute version of the Psychomotor Vigilance Performance Test (PVT) (Dinges & Powell, 1985) embedded in the AMI Motionloggers. The PVT interstimulus interval (ISI), defined as the period between the last response and the appearance of the next stimulus, ranged randomly from 2 to 10 seconds.

Procedure

This study was longitudinal and quasi-experimental in nature (non-randomized participants performing their duties in their working environment for a period of two weeks). The study protocol was approved by the Naval Postgraduate School Institutional Review Board. Personnel wishing to volunteer signed consent forms and were issued equipment for the study. Participants completed pre- and post-study questionnaires. All participants were instructed to wear their activwatch for the entire data collection period, fill out their activity logs daily and, at a minimum, complete a PVT prior to and after their watchstanding period.

Analysis

Statistical analysis was conducted with a statistical software package (JMP Pro 12; SAS Institute; Cary, NC). After assessing and rejecting the data for normality with the Shapiro-Wilk W test, comparisons were based on nonparametric methods. The Wilcoxon Rank Sum test was used for pairwise comparisons. Statistical significance in multiple comparisons was assessed using the Benjamini–Hochberg False Discovery Rate (BH-FDR) controlling procedure (Benjamini & Hochberg, 1995) with q=0.20. Correlation analysis was performed using Spearman’s rho. The criterion for statistical significance was set at p = 0.05. Data are presented as mean (M) ± standard deviation (SD).

Sleep

The average PSQI Global score at the beginning of the study was 7.13 ± 3.36, ranging from 1 to 13. PSQI scores indicated that 61.9% of the participants were “poor sleepers” (PSQI score>5). Participants working on 12-hour shifts had higher (worse) PSQI scores (8.59 ± 2.97) compared to their peers working on 8-hour shifts (6.33 ± 3.32; Wilcoxon Rank Sum test, Z=2.42, p=0.016). Consequently, 86.4% of our participants working on 12-hour shifts were identified as poor sleepers compared to 48.8% of the participants working on 8-hour shifts (Fisher’s Exact test, p=0.006).

During the entire period of the study, participants slept an average 6.74 ± 0.8 hours daily. There were no statistically significant differences in the amount of sleep received by the 8 or 12-hour shift groups (12-hour shifts: 6.84 ± 0.74 hours; 8-hour shifts: 6.76 ± 0.87 hours; Wilcoxon Rank Sums test, p=0.747). Approximately 19% (n=12) of the participants slept on average less then 6 hours per day, significantly less than the 8 hours recommended by sleep experts. Overall, they had 1.32 ± 0.26 sleep episodes in each 24-hour period. In other words, they took advantage of opportunities to nap.

Next, we assessed average daytime sleepiness. The average ESS score at the beginning of the study was 7.49 ± 3.97 improving to 6.70 ± 3.78 at the end. No statistically significant differences between watchstanding schedules were seen in the change of ESS scores (Wilcoxon Rank Sum test, Z=0.900, p=0.368).
Overall, the average Insomnia Severity Index (ISI) score at the beginning of the study was 9.36 ± 5.62, and 8.45 ± 5.04 at the end (p>0.10). Based on the ISI scores, 60% of the participants reported insomnia symptoms at the beginning of the study (45% were classified as having subthreshold insomnia, and 15% with symptoms of clinical insomnia). At the end of the study, 64% of the participants were classified as having insomnia (52% with subthreshold insomnia, 12% with symptoms of clinical insomnia). No statistically significant differences between watchstanding schedules were identified in the change in ISI scores from the beginning to the end of the study (Wilcoxon Rank Sum test, Z<0.001, p>0.990).

Consumption of Caffeinated Beverages

Approximately 70% of the participants indicated drinking caffeinated beverages, with coffee being the most frequent, followed by soft drinks (29%), and energy drinks (24%). Even though not statistically significant, those participants who followed by soft drinks (29%), and energy drinks (24%). Even though not statistically significant, those participants who reported receiving inadequate sleep also tended to report drinking caffeinated beverages at a higher rate than participants with adequate sleep.

Mood States

At the beginning of the study, POMS Total Mood Disorder (TMD) and POMS sub-scale scores did not differ between the watchstanding schedules (Wilcoxon Rank Sum test, p>0.350). The average POMS TMD score at the beginning of the study was 19.8 ± 29.6 which improved to 12.6 ± 22.1 at the end of the study period. All POMS scales improved during the course of the study; however, this trend did not differ between watchstanding schedules (Wilcoxon Rank Sum test, p>0.300).

Psychomotor Vigilance Performance

A comparison failed to identify statistically significant differences between schedules (Wilcoxon Rank Sum test, p>0.20). This result is attributed to characteristics of the 12-hour shift group, i.e., the small number of participants with useful PVT data working on 12-hour shifts (n=9), and the large differences in PVT metrics in the 12-hr shift group. The present PVT analysis, however, reveals two trends (Table 2). First, participants working on 8-hour shifts have 18% to 34% fewer errors (lapses combined with false starts) compared to their peers working on 12-hour shifts. Furthermore, the performance in 12-hour shifts is more variable and is associated with increased range of scores in all PVT metrics compared to the performance inn 8-hr shifts.

Satisfaction with Watchstanding Schedule

To assess satisfaction with the watchstanding schedules, we analyzed participant responses in three open-ended questions. First, participants responded to the question “What did you like most about your current watch schedule?” The most frequent response (n=17) was that the ADSMs liked the stability and the consistency of their schedule; they emphasized the benefit of rotating every month instead of more frequent rotations. Most of these responses were provided by ADSMs working on 8-hour shifts. Thirteen participants reported preferring 8-hour shifts. It is notable that five of the six ADSMs who rotated from 12-hour to 8-hours shifts reported preferring the 8-hour shifts.

Next, participants answered the question “What advice would you give to others who would like to improve their watchstanding schedules?” The most frequent response (n=17) was that the ADSMs who rotated from 12-hour to 8-hour shifts) was that the 12-hour watch schedule should be avoided because shifts are too long (ADSMs get tired) and the schedule does not allow for free time. In contrast to the positive responses on the 8-hour shifts, nine participants complained about the irregularity of their 8-hour watch standing schedule. Four participants working 8-hour shifts also identified that the management of work schedules and duties/activities outside watchstanding are issues of concern.

Lastly, participants answered the question, “What advice would you give to others who would like to improve their watchstanding schedules?” The most frequent response (n=17, working on 12-hour and 8-hour shifts) focused on the need for a stable schedule with a rotation of more than two weeks (many participants preferred the one-month rotation). The next most frequent response (n=10) noted the need for better management of work schedules.

DISCUSSION

Our results show that military members in our sample slept 6.74 hours on a daily basis with 19% of them sleeping on average less than 6 hours. This is less than the recommended 7 or more hours per night on a regular basis that adults aged 18 to 60 years should sleep to promote optimal health (Watson et al., 2015). Sleeping less than 7 hours per night on a regular basis is associated with adverse short term effects (impaired
performance, and greater risk of accidents), and long term effects, i.e., adverse health effects including weight gain and obesity, diabetes, hypertension, heart disease and stroke, depression, and increased risk of death (Watson et al., 2015). Results also showed that approximately 62% of the ADMSs who participated in the study were classified as poor sleepers, with 60% to 64% showing symptoms of insomnia.

We also assessed the utility of the 8-hour compared to the 12-hour watchstanding schedule. Given the complexity and the number of parameters that must be considered in assessing the utility of a watch system, Miller (2006) proposed nine “principles” which describe essential qualities of shift systems. These nine principles can be classified into three groups: circadian stability, principle of chronohygiene (short shift length, minimum number of consecutive night shifts, recovery after each night shift, maximum number of free days on weekends, at least 104 days off per year), and principles of satisfaction (equity among shift workers for types of work dates and free days, predictability of specific work and free days, and quality of time off). Based on Miller’s conceptual construct, we used a multidimensional approach to compare the 8-hour and the 12-hour schedules in terms of sleep-related factors, psychomotor vigilance performance, mood, reported preference, and the ease with which the schedule can be used in the work environment under focus.

The comparison between the two watchstanding schedules showed that military personnel working on the 8-hours shifts reported better sleep quality compared to their peers working on 12-hour shifts. Even though we did not identify statistically significant differences in psychomotor vigilance performance between the two schedules, the PVT analysis, revealed two notable trends. First, participants working on 8-hour shifts have 18% to 34% fewer errors (i.e., lapses combined with false starts) compared to their peers working 12-hour shifts. Furthermore, 12-hours shifts are associated with increased variability (i.e., a larger range of scores) in all PVT metrics compared to the 8-hr shifts. These PVT results suggest that the 8-hour watch standing schedule may be better than the 12-hour one in terms in psychomotor vigilance performance.

The open-ended responses provided insights regarding two issues of interest: which schedule the participants prefer, and the utility of the watchstanding schedules and the problems encountered when using these schedules in the operational environment. Compared to the longer 12-hours shifts, more ADMSs prefer the 8-hour shifts because they are shorter and therefore, less fatiguing. This preference for shorter shifts is supported by scientific research. Long duration shifts have been associated with elevated fatigue levels (Åkerstedt & Wright, 2009). Another study assessed the health effects of implementing a 12-hour shift in place of the traditional 8-hour shift in factory workers in Japan (Yamada, Tachibana, & Kuriyama, 1988). Compared to the workers on 8-hour shifts, workers changing to 12-hour shifts had significant increases in psychological symptoms related to fatigue and experienced, on average, a weight gain of one kilogram over a one-year period of time.

Participants also reported preferring schedules that provide regularity in work and sleep periods with less frequent rotations. But at the same time, some participants complained about the irregularity of their 8-hour schedule. In conjunction with their watch schedule, ADMSs also noted that the management of work schedules and duties/activities beyond watchstanding are issues of concern. Specifically, participants focused on the need for better management of work schedules. These results can be explained if we consider the operational environment in which our participants worked. During a typical workday, ADMSs stand watch and work on other scheduled duties/activities. ADMSs also have to respond to unscheduled operational commitments. From an ergonomics perspective, therefore, it is the combination of the load of planned activities (watch or other) and unplanned events/operational commitments with a physiologically sound watch standing schedule it is the major issue of concern. Specifically, the effective application of the watchstanding/work schedule, and the ease with which a schedule can change when needed (as is frequently the case in the dynamic environments in which our participants work).

The comparison between the two watchstanding schedules under focus provided inconclusive results in terms of sleep duration, daytime sleepiness, severity of insomnia symptoms, and mood states. The reason for this non-finding may be our participants’ actual work schedule. Even though our participants worked in rotating 8-hour and 12-hour schedules, operational commitments often interrupted the regularity of the original schedules. In general, these changes increase the variability of the obtained data. Hence, important trends tend to be masked by the “noise” of the data. This situation is typical in operational studies.

The results from our study provide a mixed picture in terms of which watchstanding schedule is better. Table 3 shows the results of the comparison between the two watch standing schedules.

Table 3. Schedule comparison.

<table>
<thead>
<tr>
<th>Factor</th>
<th>8-hour shift system</th>
<th>12-hour shift system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep attributes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep quality (PSQI)</td>
<td>Better</td>
<td>Worse</td>
</tr>
<tr>
<td>Daily sleep duration</td>
<td>Inconclusive</td>
<td>Inconclusive</td>
</tr>
<tr>
<td>Daytime sleepiness (ESS)</td>
<td>Inconclusive</td>
<td>Inconclusive</td>
</tr>
<tr>
<td>Insomnia (ISI)</td>
<td>Inconclusive</td>
<td>Inconclusive</td>
</tr>
<tr>
<td>Performance (PVT)</td>
<td>Better</td>
<td>Worse</td>
</tr>
<tr>
<td>Mood (POMS)</td>
<td>Inconclusive</td>
<td>Inconclusive</td>
</tr>
<tr>
<td>Preference</td>
<td>Preferred</td>
<td>Not preferred</td>
</tr>
<tr>
<td>Ease of application in a dynamic environment</td>
<td>More difficult to be applied and changed</td>
<td>More easily applied and changed</td>
</tr>
</tbody>
</table>

In conclusion, this study shows the difficulty of implementing a specific watchstanding schedule in operational environments overloaded with unplanned, irregular events; the scheduled work duties, combined with unplanned operational duties, frequently interrupt the flow of the typical watch schedule. Even though not conclusive, our results suggest that the 8-hour schedule is better than the 12-hour schedule in
terms of sleep and performance but may be more difficult to be applied.

Based on our results, the researchers provided to the leadership of the military organization the following recommendations to be considered:

- Use the 3-section, 8-hour watch standing schedule. If due to limited manning, a 3-section schedule is not feasible, work on a 2-section 12-hour schedule. When deemed feasible from a manning perspective, the 3-section, 8-hour schedule should be preferred.

- Three-week, or longer, watch rotations should be preferred. Frequent watch standing schedule rotations are associated with poor sleep (Arendt, Middleton, Williams, Francis, & Luke, 2006).

- Team leaders should be instructed/allowed to choose the best time shifts are changing based the operational needs, and factors specific to their work environments. For example, commute times and heavy traffic periods should be taken into account when ADSMs are required to arrive to the military facility for their morning shift (Shattuck, Matsangas, Eriksen, et al., 2015). An early morning rise will not allow the ADSMs to receive a long and recuperative night sleep.

- If operational commitments do not allow for adequate/uninterrupted sleep time, team leaders should allow for an adequate rest period after the operationally busy period.

**Study limitations**

This study had a number of limitations. The schedules to which our participants were assigned turned out to be notional rather than actual schedules due to various operational commitments. Furthermore, compliance with the PVT tests was less than we had planned – again probably due to high operational tempo and competing demands. Only nine participants on the 12-hour shifts had useful PVT data, and there was significant variability in their performance.

**Disclaimer**

The views expressed in this study are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government.

**LIST OF REFERENCES**


**LIST OF REFERENCES**


