PRIOR NIGHT SLEEP DURATION IS ASSOCIATED WITH PSYCHOMOTOR VIGILANCE IN A HEALTHY SAMPLE OF POLICE ACADEMY RECRUITS

Thomas C. Neylan,1,2,3 Thomas J. Metzler,1,2,3 Clare Henn-Haase,1,2,3,5 Yelena Blank,2,4 Gary Tarasovsky,2,3 Shannon E. McCaslin,1,2,3 Maryann Lenoci,2,3 and Charles R. Marmar1,2,3,5

1Department of Psychiatry, University of California, San Francisco, San Francisco, California, USA
2Northern California Institute for Research and Education, San Francisco, California, USA
3Mental Health Service, San Francisco Department of Veterans Affairs Medical Center, San Francisco, USA
4Department of Psychology, University of Arizona, Tucson, Arizona, USA
5Department of Psychiatry, New York University School of Medicine, New York, USA

Aviation, military, police, and health care personnel have been particularly interested in the operational impact of sleep restriction and work schedules given the potential severe consequences of making fatigue-related errors. Most studies examining the impact of sleep loss or circadian manipulations have been conducted in controlled laboratory settings using small sample sizes. This study examined whether the relationship between prior night sleep duration and performance on the psychomotor vigilance task could be reliably detected in a field study of healthy police academy recruits. Subjects (N = 189) were medically and psychiatrically healthy. Sleep-wake activity was assessed with wrist actigraphy for 7 days. Subjects performed the psychomotor vigilance task (PVT) for 5 min on a personal digital assistant (PDA) device before and after their police academy workday and on comparable times during their days off. Mixed-effects logistic regression was used to estimate the probability of having ≥1 lapse on the PVT as a function of the previous night sleep duration during the 7 days of field testing. Valid estimates of sleep duration were obtained for 1082 nights of sleep. The probability of a lapse decreased by 3.5%/h sleep the night prior to testing. The overall probability of having a lapse decreased by 0.9%/h since awakening, holding hours of sleep constant. Perceived stress was not associated with sleep duration or probability of performance lapse. These findings demonstrate the feasibility of detecting sleep and circadian effects on cognitive performance in large samples.
field studies. These findings have implications regarding the daytime functioning of police officers. (Author correspondence: Thomas.Neylan@ucsf.edu)

**Keywords**  Circadian rhythm/physiology; Cognitive performance; Intrinsic alertness; Police; Reaction time; Sleep deprivation/complications; Vigilance

**INTRODUCTION**

There is increasing public awareness that sleep duration has implications for a number of economic and health outcomes, particularly operational productivity, errors, and accidents (e.g., Banks & Dinges, 2007; Barger et al., 2006; Buysse et al., 2003; Dorrian et al., 2006; Landrigan et al., 2004; Lockley et al., 2007; Philibert, 2005). Aviation, military, police, and health care personnel have been particularly interested in the operational impact of sleep restriction given the potential severe consequences of making fatigue-related errors. This field study examines the relationship between prior night sleep duration, performance on the psychomotor vigilance task (PVT), and perceived stress in a healthy sample of police academy recruits without current psychopathology. Data are taken from the baseline assessment of a large prospective study of risk and resilience of police academy recruits who are assessed at baseline during police academy training and again every 6 months during active duty police service. The overall goal of the prospective study is to determine if short sleep duration, shiftwork, and long work hours affect psychological resilience to exposure to routine work-related stressors and critical incident exposure in new police officers during their subsequent years of police service. The current study addresses whether the relationship between prior night sleep duration and performance on the PVT could be reliably detected in a field study of healthy police academy recruits during the baseline assessment period.

A large body of literature demonstrates that sleep restriction, conducted in controlled laboratory settings, is associated with poor intrinsic alertness (Balkin et al., 2004; Belenky et al., 2003; Carskadon & Dement, 1981; Davidson et al., 1987; Dinges et al., 1997; Friedman et al., 1977; Horne & Wilkinson, 1985; Miccoli et al., 2008; Rosekind et al., 1994; Sallinen et al., 2008; Van Dongen et al., 2003; Webb & Agnew, 1974). Intrinsic alertness is defined as the ability to prepare and maintain a readiness to respond to stimuli in the absence of a warning signal (Fimm et al., 2009; Perin et al., 2010; Sturm et al., 1999). The term psychomotor vigilance, when used to describe performance on a simple reaction time task, is synonymous with term intrinsic alertness. Generally, it has been shown that chronic partial sleep loss results in a cumulative and dose-dependent decline in multiple measures of intrinsic alertness, particularly the ability to sustain attention and respond quickly to a signal, as
measured by reaction time and performance lapses on the PVT (Belenky et al., 2003; Dinges et al., 1997; Van Dongen et al., 2003, 2004). Sleep loss has also been shown to adversely impact reaction time when measured with instruments other than the PVT (Bonnet, 1985; Bonnet & Rosa, 1987).

Greater performance deficits on the PVT have consistently been observed with experimental sleep restriction of ≤4 h sleep (e.g., Dinges et al., 1997), and sleep deprivation effects appear to be cumulative, with performance decrements increasing across successive days of sleep restriction (e.g., Dorrian et al., 2003). Further, even moderate sleep deprivation, ranging from 5 to 7 h, can negatively impact performance on the PVT (Dinges et al., 1997). Moderate sleep deprivation has been shown to lead to significantly poorer performance on accuracy measures, such as hand-eye coordination and reaction time, comparable to subjects with up to 0.1% blood alcohol concentration (Williamson & Feyer, 2000). Chronic partial sleep deprivation is also associated with neurocognitive deficits in higher-order domains, such as executive function (for review see Durmer & Dinges, 2005). Belenky et al. (2003), in a study of 66 healthy subjects in a laboratory setting, found a sleep-dose-dependent relationship with PVT performance. With a severe level of sleep restriction (3 h), performance declined consistently over 7 days with no stabilization. PVT reaction time continued to increase, and the number of lapses on the PVT increased steadily across the 7 days. However, when sleep was restricted to a moderate level of 5–7 h, an initial decrease in performance was observed, but performance stabilized at a reduced level after a few days and was sustained throughout the week. A non-sleep-restricted (9 h) subject group was also included, and the performance of these individuals remained at baseline levels.

The majority of studies examining PVT performance have a number of limitations. Most have been conducted in laboratory settings with relatively small sample sizes. Laboratory settings lack the real-life salience of operational environments and, therefore, may have greater sensitivity to detect a relationship between sleep duration and cognitive performance. Field studies are needed to demonstrate the ecological validity of the observed laboratory findings and to demonstrate that the effects of sleep duration on performance are not masked in operational environments. The development of a (personal digital assistant) PDA version of the PVT (Thorne et al., 2005) has opened up the possibility of testing the relationship between sleep duration and intrinsic alertness in large field studies. The aim of this study was to examine the relationship between an objective estimate of sleep duration with wrist actigraphy, perceived psychological stress, and intrinsic alertness on the PVT in a healthy sample of police academy recruits without current psychopathology. Our secondary aim was to test the validity of the PDA PVT measure against the gold standard
laboratory device. Our hypothesis was that in a naturalistic setting of healthy subjects, shorter mean sleep duration would be associated with worse performance on psychomotor vigilance testing.

MATERIALS AND METHODS

Participants and Procedures

Participants were recruited from four urban police departments, the New York Police Department (NYPD) and three departments in the San Francisco Bay Area (Oakland, OPD; San Francisco, SFPD, and San Jose, SJPD) during police academy training. The study was approved by the University of California San Francisco Committee on Human Research and the Clinical Research Review Committee of the San Francisco VA Medical Center and meets the international ethics standards of the journal (Portaluppi et al., 2008). All procedures were carried out with the adequate understanding and written consent of the subjects.

Academy trainees were introduced to the study through an in-person presentation made by study personnel during academy training classes. This presentation included the distribution of two letters, one from the commissioner or police chief of the affiliated department and one from the study team. Included with the approach letters was a description of the study procedures, a contact number, and a participation form that included the option to be contacted by the study team. Recruitment also involved informational flyers posted at each academy that included contact information for the study team. Academy recruits who were combat veterans or had prior experience in law enforcement or emergency services occupations were excluded. Actigraphy and psychomotor vigilance testing were obtained in 189 subjects with a mean age of 26.4 (SD ± 4.2 and range 21–43) yrs. The sample was 88% male and medically and psychologically healthy with no current Axis I psychiatric disorder as determined by a clinical interview using the Clinician Administered PTSD Scale (CAPS) (Blake et al., 1995) and the Structured Clinical Interview for DSM-IV, version 2.0 (SCID I-NP) (First et al., 1996). None of the police academy recruits worked the night shift during this phase of their training.

Measures

Subjective sleep quality was assessed at the baseline evaluation with the Pittsburgh Sleep Quality Index (PSQI) (Buysse et al., 1989). The PSQI is a 19-item self-report questionnaire, previously validated by polysomnography, that assesses sleep quality and disturbances that have occurred during the past month. The 19 individual items yield seven component scores:
subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction. Scores on the seven dimensions are summed to form a single global score. A PSQI score >5 has been shown to have sensitivity of 89.6% and specificity of 86.5%, with kappa = 0.75 for distinguishing good sleepers from clinical insomnia subjects (Buysse et al., 1989).

Subjective stress was assessed at the baseline evaluation with the Perceived Stress Scale (PSS) (Cohen et al., 1983). The PSS is a 14-item self-report measure of nonspecific appraised level of stress in one's life during the past month. Response choices for each item uses a 0–4 Likert scale. The PSS was shown to have adequate reliability (coefficient α = .84–.86; test-retest = .55–.85) and adequate concurrent and predictive validity (Cohen et al., 1983).

Sleep-wake activity was assessed by wrist actigraphy (Octagonal Basic Motionlogger; Ambulatory Monitoring) for 7 consecutive days. We employed the Proportional Integrating Measure (PIM)/Zero Crossing Mode/Time Above Threshold sampling mode in 1-min epochs and used the PIM University of California San Diego (UCSD) algorithm of Action 4 software to estimate sleep duration/24-h period. The PIM UCSD algorithm has been found to have the best accuracy for estimating total sleep time in young healthy subjects (Jean-Louis et al., 2001). The primary sleep measure was total sleep duration/24-h period over 7 days. Subjects were instructed to perform the PVT for 5 min on a PDA device immediately pre- and postshift and at comparable times on days off during the same 7 days of actigraph monitoring (Thorne et al., 2005). There were no night shifts during academy training. Prior to the 7 days of monitoring, subjects came to the laboratory and completed successive 5-min PVT trials with the PDA device and the standard PVT device (PVT 192; Ambulatory Monitoring) in random counterbalanced order.

The PVT is a widely used instrument that measures sustained attention, or more precisely, intrinsic alertness, via measuring simple reaction time (Dinges & Powell, 1985). Extensive work with this measure has demonstrated that the PVT is not affected by practice effects and is a highly sensitive measure of the effects of disrupted circadian rhythms from shiftwork (Dorrian et al., 2003; Rosekind et al., 1994) and chronic sleep deprivation (Belenky et al., 2003; Van Dongen et al., 2003). The PVT has a random interstimulus interval of 2–10 s and can be collected over a preset time period. The 5-min PVT task was chosen for pragmatic reasons to limit subject burden and improve adherence to the study protocol. The 5-min PVT has been found to be strongly correlated to performance on the standard 10-min task and a useful alternative when response burden is a concern (Loh et al., 2004; Roach et al., 2006). Subjects were given an incentive, a modest monetary bonus, to complete the PDA and other study tests. The main measure used in this study was
the number of performance lapses, i.e., reaction time >500 ms/5-min period. Each data entry was digitally time stamped to monitor adherence and to test for effects of time since awakening. Custom made PVT software (Ambulatory Monitoring) for PDA devices was modified from the PVT software developed by the US Army (Thorne et al., 2005). PDA PVT uses a pseudorandom interstimulus interval and a bulls-eye target, in contrast to the PVT 192, which uses a bright red-light stimulus. Prior to the week of testing, subjects were seen for a clinical assessment and were trained on the use of the PVT. After training, the subjects performed the PVT with both the PDA device (5 min) and the gold standard PVT 192 Psychomotor Vigilance Task Monitor (5 min) in a randomized counterbalanced order in order to conduct a validity check on the PDA PVT device.

**Statistical Analyses**

Mixed-effects logistic regression for repeated-measures was used to estimate the probability of having ≥1 or more lapse (>500 ms reaction time) on the PVT as a function of sleep duration the previous night, and lag time between sleep offset and the time of the PVT trial. The dependent variable was the dichotomized number of lapses in each PVT trial (0 versus ≥1 lapses). The mixed model includes a random intercept to accommodate within-subject correlations among the responses, using the xtmelogit command in Stata version 10 software (StataCorp, 2007). Advantages of the mixed model are that it models both within- and between-subject variations and properly accounts for within-subject repeated-measures in calculating standard errors and significance tests. It also allows one to use all available data from each subject, without losing data to case-wise deletion as with repeated-measures analysis of variance (ANOVA) (e.g., Brown & Prescott, 1999). The primary predictors in the model were number of hours of sleep the night before a given PVT trial and the number of hours from awakening to a given trial. Squared terms were added for both hours of sleep and hours from awakening to assess curvilinear effects of these variables. Clock time was accounted for using a sine transformation as suggested by Van Dongen et al. (2007). The phase angle was adjusted to achieve a nadir at 05:00 h, as suggested by Achermann and Borbely (1994). Although the probability of a lapse was chosen a priori as our primary outcome, median reaction times were also analyzed, as well as mean reaction times for the slowest and fastest 10% of responses for each trial using linear mixed-effects models. As a further check on the robustness of the findings, the raw number of lapses in each trial (ranging from 0 to 31) was modeled using a random effects negative binomial model for count data (Stata command “xtnbreg”).
RESULTS

Subjective sleep quality in the month prior to the baseline assessment as measured by the PSQI shows a mean global score of 3.5 (SD ± 2.1). This is within the range of healthy sleepers (Buysse et al., 1989) and substantially lower than the mean global PSQI score of 6.2 in our prior active duty police sample (Neylan et al., 2002).

Valid actigraph estimates of sleep duration were obtained for 1082 nights of sleep (82%) out of a possible total of 1323 nights. The primary reasons for data loss were battery failure, loss of electrical connectivity to the battery, or subjects removing the actigraph and forgetting to replace it before sleep onset. Only 1.3% of nights of sleep were followed by a nap period. Mean duration of total sleep/24 h was 6.25 h (SD ± 1.70). Mean sleep duration not including naps was 6.24 h (SD ± 1.71), demonstrating the negligible presence of naps in this sample. A frequency histogram of the distribution of total sleep time/24 h is presented in Figure 1. The mean wake time was 7.8 h after midnight (SD ± 2.9). The mean onset of sleep time was 1.6 h (SD ± 2.5) after midnight.

PVT Validation Study

Complete data on successive PVT tasks comparing the PVT 192 to the PDA PVT were obtained in 143 subjects. The order of PVT 192 to PDA PVT was randomized. Mean reaction time was longer on the PDA PVT (270 ms, SD ± 54) compared to the gold standard PVT 192 device (246 ms, SD ± 36; t(142) = 6.16, p < .001). Similarly, median reaction time on the PDA device was longer than on the PVT 192 (252 ms, SD ± 38,
compared to 231 ms, SD ± 28; \( t(142) = 9.33, p < .001 \). The correlation between mean reaction times on the two devices was \( r = .54 \), whereas the correlation for medians was \( r = .72 \), both \( p < .001 \). The mean number of lapses per trial was .89 (SD ± 2.2) on the PDA PVT compared to .55 (SD ± .2) on the PVT 192; \( t(142) = 2.45, p < .05 \). The correlation of PDA PVT and PVT 192 lapses was high (\( r = .68, p < .001 \)). This is not as high as the published test-reliability for PVT 192 for both lapses (intraclass correlation coefficient [ICC] = .88, \( p < .0001 \)) and median reaction time (ICC = .83) (Van Dongen et al., 2003).

### 7-Day Palm PVT Results

A total of 1925 PVT trials following a valid estimate of preceding sleep was obtained. The mean reaction time (RT) (mean of the mean RT for each trial) was 299 ms (SD ± 137); the mean of all subjects’ median response times was 269 ms (SD ± 53); the mean of the fastest 10% RTs was 216 ms (SD ± 30); and the mean of the slowest 10% RTs was 554 ms (SD ± 827). The mean number of false starts was 1.5 (SD ± 5.1). Only 9 of the 1925 PVT trials (<0.5%) in our study had any nonresponses/30 s) for one or more targets. Nonresponses were not counted as lapses. All nine of these trials had good responses to analyze. Of the nine, seven had at least one lapse, and two had no lapses. 53.7% of PVT trials showed \( \geq 1 \) lapse. Frequency histogram of the distribution of the number of PVT lapses is presented in Figure 2. The distribution of the interval in hours between time of sleep offset for all PVT trials pre- and postshift (lag time) is presented in the frequency histogram in Figure 3.

![FIGURE 2](image)

**FIGURE 2** Frequency histogram of the distribution of the number of PVT lapses in 1925 PVT trials in 189 healthy police academy recruits.
The mixed-effects logistic regression model predicting an occurrence of ≥1 lapse is presented in Table 1. The primary predictors were hours of sleep the previous night, time in hours since awakening, and squared terms for both. The predictors were entered simultaneously, so each effect is controlled for by the others. However, the correlations among the predictors were all <.08. Sine-transformed clock time was added to the model and found to explain no additional variance (likelihood ratio $\chi^2(1) = 0.0, p = .95$) and therefore excluded. We did not find a significant interaction between prior night sleep duration and time since awakening.

The probability of having a lapse decreased by 3.5%/h of sleep the subject had the previous night. As expected, the probability was highest with short sleep duration. After 3 h of sleep, subjects had a 67% calculated probability of a lapse in contrast to 50% after 8 h, or 43% after sleeping 10 h. The overall probability of having a lapse decreased by 0.9%/h since awakening, holding hours of sleep constant at its mean. However,

### TABLE 1 Repeated-measures logistic regression model predicting ≥1 lapses on the PVT

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Odds ratio&lt;sup&gt;a&lt;/sup&gt; (95% CI)</th>
<th>Change in probability&lt;sup&gt;a&lt;/sup&gt; (95% CI)</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep (h)</td>
<td>0.87 (0.80–0.94)</td>
<td>−3.5% (−1.5% to −5.5%)</td>
<td>−3.48</td>
<td>.001</td>
</tr>
<tr>
<td>Sleep&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.01 (0.98–1.04)</td>
<td>—</td>
<td>0.75</td>
<td>.455</td>
</tr>
<tr>
<td>Hours since awakening</td>
<td>0.97 (0.95–0.98)</td>
<td>−0.9% (−0.4% to −1.3%)</td>
<td>−3.46</td>
<td>.001</td>
</tr>
<tr>
<td>Hours since awakening&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.004 (1.000–1.008)</td>
<td>—</td>
<td>1.97</td>
<td>.048</td>
</tr>
</tbody>
</table>

<sup>a</sup>Odds ratios (ORs) and changes in probability of a lapse are the effects of each additional hour of sleep or time since awakening.

<sup>b</sup>Sleep and hours since awakening represent the squared terms in the logistic regression model.
there was a significant curvilinear effect of hours since awakening, which demonstrates that the risk of lapse varies from morning to evening hours, independent of sleep duration.

Figures 4 and 5 show both observed and logistic model-predicted responses as functions of the five predictors, smoothed using the lowess smoother in Stata v. 10. The probability of having ≥1 lapses on the PVT as a function of the previous night of sleep (plus nap time if present prior to trial) is presented in the lowess plot depicted in Figure 4. The probability of having ≥1 lapse on the PVT as a function of hours since awakening is presented in the lowess plot depicted in Figure 5. The significant curvilinear effect shows that the risk for lapse accelerates during the last period of the wake period.

Results for the linear mixed-effects model using median RT as the dependent variable showed similar results. Median RT decreased with longer sleep duration (β = −1.5, z = 2.35, p < .05), and there was a trend of its reduction as a function of hours since awakening (linear term β = .28, z = 1.77, p = .07), until the end of the wake period when it increased in a curvilinear function (squared term β = .21, z = 4.30, p < .001). Additionally, the mixed-effects negative binomial regression showed the same pattern of results using total number of lapses as the outcome (incidence rate ratio [IRR] for sleep duration = .94, z = 3.26, p < .001; for hours awake linear term IRR = .98, z = 3.00, p < .01; for hours awake squared term IRR = 1.00, z = 1.63, p = .10). Analyses of the log-transformed means of the slowest and fastest 10% of responses showed an effect of sleep duration on the slowest responses (β = −.08, z = 3.53, p < .001), but not the fastest responses (β = −0.01, z = 1.66, p = .10). Hours awake affected both the fastest responses (linear term beta = .00, z = 1.06,
$p = .29$; squared term $\beta = .16$, $z = 10.95$, $p < .001$) and the slowest responses (linear term $\beta = .01$, $z = 2.55$, $p < .05$; squared term $\beta = .16$, $z = 2.80$, $p < .01$). Note that the effects of hours awake are curvilinear, and reflect increasingly slow reaction times towards the end of the period of wakefulness.

Perceived stress was not associated with sleep duration ($r = -.07$, $p = .11$) or probability of performance lapse.

**DISCUSSION**

In a healthy sample of subjects with relatively short mean sleep duration, the probability of having $\geq 1$ performance lapse on the PVT was inversely associated with the length of the previous night of sleep. Median RT also decreased with longer duration of prior night sleep. This is consistent with the observed dose-response relationship between sleep duration and performance on the PVT observed in controlled laboratory conditions (Belenky et al., 2003; Jewett et al., 1999; Van Dongen et al., 2003) and a semilaboratory environment (Axelsson et al., 2008). These findings demonstrate the feasibility of measuring sleep and intrinsic alertness in large field studies as well as the robust relationship between sleep duration and performance on the PVT. Further, these findings have clear implications regarding the daytime functioning of police officers. Shorter sleep duration is associated with diminished ability to remain attentive and respond in a timely manner. Considering that short sleep duration negatively impacted performance in this sample of new recruits who, although in a stressful training setting, have not yet been exposed to critical incident stress nor experienced variables such as shiftwork,
highlights the importance of early attention to this aspect of psychological health. This might include providing education and tools for promoting healthy sleep even while receiving training in the academy.

Median RT and mean number of lapses were slightly, but significantly, higher on the PDA PVT as compared to the gold standard PVT 192 device. This is similar to results reported by Lamond et al. (2008). These data could be viewed as an argument both for and against the use of the PDA device. Overall, the PDA was highly correlated with the gold standard PVT 192 device and the observed relationship between sleep duration and probability of lapse argues for the validity of the PDA device. However, the longer reaction time and higher frequency of lapses with the PDA compared to the standard device, measured in immediate succession in a counterbalanced order, suggests that there is mechanical biasing in the PDA to inflate the measure of lapses. Nevertheless, given the ease with which our subjects were able to carry and manage the PDA device, we can conclude that this method has high operational utility in large field studies. The standard PVT might remain the optimal choice for field studies where it is feasible for participants to conduct the testing in the same setting. For example, Gander et al. (2008) conducted a field study of anesthesiologists who used a quiet hospital office for PVT testing. They were able to detect a progressive decline in performance on the PVT, conducted without supervision by study personnel, in anesthesiologists who worked 12 consecutive days and averaged a restriction in sleep time <1 h/ day.

The raw data examining observed sleep time and the probability of a performance lapse would suggest that improvement in intrinsic alertness levels off above a particular threshold of sleep duration (e.g., 8 h), as was found in another study (Jewett et al., 1999). However, the calculated probability of lapse that also accounts for time since awakening shows that intrinsic alertness continues to improve in the interval from 8 to 10 h of sleep. The relatively short mean sleep duration of this cohort suggests that our sample is reflective of the high prevalence of chronic partial sleep loss in the United States (Kripke et al., 2002; National Sleep Foundation, 2008). It is possible that our near-linear relationship between prior sleep duration and probability of performance lapse is explained by having a sample that is chronically sleep restricted and thus nights of sleep >8 h continue to produce benefits.

Within individual trials each day, there was a weak, but significant, decrease in the probability of a lapse as a function of time since awakening. However, the significant curvilinear effect demonstrates that this can only be interpreted when considering the point on the curve when the probability is calculated. The predicted and observed curves demonstrate a slight decrease in risk across the day, consistent with known circadian alerting effects, and an abrupt increase in lapses in the terminal part of
the day prior to sleep initiation. Similarly, median RT decreased across the day until the terminal part of the wake span when it abruptly increased. This is consistent with other studies that have demonstrated the sensitivity of the PVT to circadian-related variability in performance when measured in controlled laboratory settings. The replication of this finding in a naturalistic field study using an inexpensive PDA device demonstrates the robust effect of prior sleep duration and time since awakening on risk for a performance lapse. We conclude that we are able to detect effects of circadian-related variability in intrinsic alertness on the PVT, as has been shown with studies using the standard laboratory device (Dinges et al., 1997; Graw et al., 2004).

The study has several limitations. The first relates to the use of actigraph estimates of sleep duration, which are not as accurate as the gold standard polysomnogram. However, actigraph estimates of sleep times have been shown to be highly accurate in healthy subjects (Ancoli-Israel et al., 2003). The second relates to a common issue of all field studies; that is, the psychomotor vigilance testing was conducted by subjects in their natural environment under variable circumstances. We gave careful instructions to our subjects regarding the importance of finding a quiet setting to conduct the PVT test, and subjects had a monetary incentive to perform well. However, we have no direct way of determining what sources of distraction were present at the time of testing. The predominantly young male sample limits the ability to examine possible age and sex effects on sleep and performance (Monk et al., 2009; Tonetti et al., 2008). Finally, performance on the PVT may not be an adequate proxy for real world operational performance.

Our finding of a relatively short mean sleep duration in our subjects is comparable to data in national surveys, which have found a widespread prevalence of voluntary sleep restriction (e.g., Basner et al., 2007; Russell et al., 2007). Our sample was not selected to be representative of all police academy recruits. However, the relatively short sleep duration of our police academy subjects was somewhat unexpected and worthy of comment. We are collecting prospective data in this cohort to test whether entry into active duty police work will further degrade sleep quantity and quality given the stress of critical incident exposure and the demands of rotating and overtime work shifts. This follows from data from a number of surveys in police departments across the United States that demonstrate that a third of police officers work ≥20 h of overtime/month and that over half moonlight (Vila & Kenney, 2002). Further, mean sleep duration of <6.5 h and a high prevalence of insomnia has been found in several studies of active duty police officers (Neylan et al., 2002; Vila, 2000). Controlled restriction of sleep duration to ≤6 h over a period of 2 wks produced performance lapses on the PVT equal to two nights of total sleep loss (Van Dongen et al., 2003). The short sleep
duration of our academy recruits suggests that our subjects act much like trainees in multiple other professions, in that they do not wait for the additional work demands to unfold before voluntarily restricting their sleep time. Rather, like many in society, they make choices about balancing their time at work with additional activities to optimize their wake time while experiencing as little of the impact of short sleep as possible. The lack of an association between sleep duration and perceived psychological stress in this healthy sample suggests that the short average sleep time is a product of choice and not a marker of distress. Unfortunately, the ability to self-monitor the cognitive effects of sleep restriction is not accurate or reliable (e.g., Dorrian et al., 2003; Van Dongen et al., 2003). The increased risk of performance lapse may be particularly problematic for police officers during operations such as driving or stakeouts that require sustained vigilance.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (C.R.M.: MH56350; T.C.N.: MH73978).

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES


StataCorp. (2007). Stata statistical software, release 10. College Station, TX: StataCorp.


