There is no doubt that sleep is important for cognitive performance. Although the functions of sleep are not yet fully understood, its relationship to performance is evident through the deterioration of cognitive functioning under conditions of sleep deprivation and the recuperation provided by subsequent sleep [1,2]. The alternation of sleep and wakefulness is driven by a complex neurobiology that has only partially been unraveled [3]. Nevertheless, two primary processes of sleep/wake regulation have been putatively distinguished [4].

The first process, referred to as the sleep homeostat, seeks to balance time spent awake and time spent asleep. It can be conceptualized as the buildup of homeostatic pressure for sleep during periods of wakefulness, and the dissipation of this pressure during periods of sleep. The second process is the endogenous circadian rhythm, which is driven by the biological clock in the suprachiasmatic nuclei of the hypothalamus in the brain. This “internal clock” keeps track of the time of day (the term circadian refers to a near–24-hour cycle). Given that humans are a diurnal species, the circadian process seeks to place wakefulness during the day and sleep during the night. The circadian process can be envisaged as providing pressure for wakefulness [5,6] that is strongest during
the early evening hours and weakest in the early morning. The sleep homeostatic process and the circadian process interact with each other neurobiologically.

During the day, the homeostatic and circadian processes act in opposition to promote wakefulness [7]. In the morning hours just after awakening from a sleep period, not much homeostatic pressure for sleep is present, and there is also relatively little compensatory circadian pressure for wakefulness. (The awakening process itself is characterized by a brief period of “sleep inertia,” which is discussed later.) As the day progresses, the homeostatic pressure for sleep builds up, and at the same time, the circadian pressure for wakefulness increases. The net effect is stable waking pressure throughout the day, which in healthy individuals results in a consolidated period of wakefulness.

At night, the homeostatic and circadian processes act synergistically to promote sleep [7]. In the beginning of the night before falling asleep, the circadian pressure for wakefulness gradually withdraws, whereas the homeostatic pressure for sleep continues to accumulate. As a result, there is a notable net increase in pressure for sleep and, under appropriate circumstances (e.g., lying supine), the sleep state is initiated. During sleep, the homeostatic pressure for sleep dissipates. The circadian pressure for wakefulness further diminishes as well. Thus, there is little net waking pressure throughout the night, which in healthy individuals results in a consolidated period of sleep.

In the morning, the circadian pressure for wakefulness gradually rises again and exceeds the largely dissipated homeostatic pressure for sleep. Consequently, spontaneous awakening occurs, and the cycle starts again with the homeostatic and circadian processes acting in opposition to promote wakefulness. These interactions of sleep homeostatic and circadian neurobiology have been studied extensively [8] and have been instantiated in contemporary theoretical and mathematical models [9,10].

Even though the alternation of sleep and wakefulness is regulated fairly precisely, giving rise to the term sleep/wake cycle, humans are a rather unique species in that they can voluntarily choose to temporarily ignore the homeostatic and circadian-mediated signals for sleep [11]. When humans stay awake to pursue other activities, though, this is not without consequence, which will be discussed later.

**Homeostatic and circadian influences on performance**

To systematically examine the effects of the homeostatic pressure for sleep and the circadian pressure for wakefulness on cognitive performance, studies have been conducted in laboratories set up specifically to monitor and control sleep and wakefulness, circadian rhythms, and neurobehavioral functions [1,12,13]. This article focuses on psychomotor vigilance performance, because it involves reaction time and sustained attention, which are elemental features of a wide range of human performance.
Psychomotor vigilance performance can be measured with the psychomotor vigilance task (PVT) [14], a portable, easily usable reaction-time test with a high stimulus load (visual or auditory) that can yield rapid (ie, in 10 minutes) and reliable assessments of psychomotor vigilance impairment [15,16]. The PVT has been used in the laboratory to precisely measure, at brief intervals (typically every 2 hours of wakefulness), the changes in psychomotor vigilance performance caused by sleep loss and circadian rhythmicity [16].

The changes in psychomotor vigilance performance over time in a laboratory study involving 88 hours of extended wakefulness (ie, three nights without sleep) [15] are shown in Fig. 1. In Fig. 1A, lapses of attention (ie, reaction

Fig. 1. Performance data from 13 healthy young adult males (mean age ± SD: 27.3 ± 4.6 years) who spent 10 days in the controlled environment of a laboratory. After one adaptation day and two baseline days with 8 hours time in bed (23:30–07:30), they were assigned to a condition involving 88 hours of extended wakefulness. Thereafter, during the last 3 days of the experiment, they received recovery sleep each night. A subset of 7 subjects were allowed 7 hours time in bed (23:30–06:30) on the first 2 recovery days and 14 hours time in bed (23:30–13:30) on the last recovery day, whereas the other 6 subjects were allowed 14 hours time in bed on all 3 recovery days. Throughout scheduled wakefulness, subjects underwent cognitive testing every 2 hours. The cognitive test battery included a 10-minute psychomotor vigilance task (PVT). (A) The number of lapses (reaction times ≥ 500 ms) on the PVT. (B) The average of the 10% fastest reaction times (in ms) on the PVT. In both cases, group averages are plotted against cumulative clock time. Gray bars indicate scheduled sleep periods—the 2 baseline nights and the first 2 recovery nights (7 hours time in bed) are shown. Dotted lines in the 88-hour sleep deprivation period indicate midnight. On the last baseline day (before the last baseline sleep period) and on the first day of sleep deprivation, psychomotor vigilance lapses were relatively rare and fastest reaction times were relatively short. However, both psychomotor vigilance lapses and fastest reaction times increased significantly during the rest of the 88 hours of wakefulness. The progressive increases over days of sleep deprivation were modulated by a circadian rhythm: the number of lapses and the fastest reaction times were reduced during the diurnal hours compared with the nocturnal hours even after 3 days without sleep. Recovery sleep rapidly reduced the level of impairment; after 2 nights with 7 hours in bed, performance was almost back to the baseline level (for the recovery days, averages are shown for the 7 subjects who received 7 hours time in bed only).
times $\geq 500$ milliseconds) on the PVT can be seen, whereas Fig. 1B shows the 10% fastest reaction times on the PVT. The changes in fastest reaction times showed the same temporal pattern as the changes in psychomotor vigilance lapses. With every day of sleep deprivation, the average number of psychomotor vigilance lapses and the average duration of the fastest reaction times increased. In addition, nighttime performance was consistently worse than daytime performance.

This temporal pattern can be readily interpreted as the interaction between the homeostatic and circadian processes of sleep/wake regulation. The two processes are considered to have a combined effect on waking cognitive performance, which can be approximated by subtracting the homeostatic pressure for sleep from the circadian pressure for wakefulness [17]. The net pressure for sleep determines the degree of cognitive performance impairment. This explains why in the study of Fig. 1, psychomotor vigilance lapses and fastest reaction times increased over days of sleep deprivation, since the homeostatic pressure for sleep continued to accumulate in the absence of sleep. Furthermore, it explains why daytime performance was consistently better than nighttime performance, for the circadian pressure for wakefulness was greatest during the diurnal portion of each day.

The interaction of the homeostatic and circadian processes can also be observed in the phenomenon of jet lag. This is the transient period of impairment following rapid travel to a different time zone. On arrival in the new time zone, the circadian pressure for wakefulness is initially not timed appropriately relative to the time of day, which has an adverse effect on daytime cognitive performance. The circadian pressure for wakefulness is also not withdrawn at the right time to promote nighttime sleep. This may cause problems with the timing and consolidation of sleep, resulting in reduced dissipation of the ho-
meostatic pressure for sleep. The remaining pressure for sleep may further compromise cognitive performance. Depending on the direction of travel and the number of time zones crossed, it can take several days for the circadian process to align properly with the new time zone and for the homeostatic process to restore the balance between sleep and wakefulness [18].

Effects of napping on performance

The dissipation of homeostatic pressure during sleep is thought to be an exponential process [9], where the greater the level of homeostatic pressure reached during wakefulness, the faster the dissipation during subsequent sleep. This implies that recuperation from performance deficits caused by prior sleep loss should occur rapidly even if time available for sleep is relatively short. The data in Fig. 1 (see right-hand side of Figs. 1A and 1B), which show that a single episode of 7 hours time in bed markedly reduced psychomotor vigilance lapses and fastest reaction times after the 88 hours of total sleep deprivation, confirm this.

Based on the disproportionate recovery potential of relatively short sleep periods, naps (“power naps”) have been investigated as a strategy to attenuate performance deficits during and following periods of sleep deprivation [19,20]. Fig. 2 illustrates the effects of nap sleep on psychomotor vigilance lapses during

Fig. 2. Performance data from 13 healthy young adult males (mean age ± SD: 28.2 ± 8.9 years) who participated in the same experiment as those of Fig. 1, but were randomized to a condition involving 2-hour nap opportunities every 12 hours (14:45–16:45 and 02:45–04:45) during 88 hours of otherwise continuous wakefulness. Details of the figure are the same as for Fig. 1A, but there is a difference in the range of the ordinate scale. Thin gray bars indicate the scheduled 2-hour nap periods. A subset of 8 subjects were allowed 7 hours time in bed on the first 2 recovery days, whereas the other 5 subjects were allowed 14 hours time in bed. For these recovery days, averages are shown for the 8 subjects who received 7 hours time in bed only. Compared with 88 hours of total sleep deprivation (Fig. 1A), the 2-hour nap opportunities considerably attenuated the magnitude of psychomotor vigilance deficits from sleep loss, although a minor buildup of performance impairment still occurred across the 88-hour experimental period. Nap sleep resulted in vigilance performance deficits immediately on awakening, however. This “sleep inertia” effect intensified with progressive sleep loss, especially at night.
otherwise continuous wakefulness. The experimental condition depicted in this figure was comparable to the 88 hours of extended wakefulness shown in Fig. 1, but in this case the 88 hours were interrupted by 2-hour nap opportunities occurring every 12 hours [15,21]. As a consequence, the buildup of psychomotor vigilance impairment over the 88-hour period was considerably attenuated, which highlights the recuperative potential of nap sleep.

Unfortunately, the napping strategy has an adverse effect called *sleep inertia*, which is the cognitive performance impairment commonly experienced immediately after awakening [22]. As evident in Fig. 2, sleep inertia is particularly noticeable under conditions of sleep loss and during the circadian night [23–25]. Thus, the magnitude of sleep inertia appears to be a function of increased homeostatic pressure for sleep and decreased circadian pressure for wakefulness. In situations where optimal performance capability right after awakening is not required, napping may still be useful to mitigate the effects of sleep loss. Also, very short naps (approximately 10 minutes) may offer some recuperative benefit without noticeable levels of sleep inertia [26]. Even so, strategic napping cannot be considered a universal substitute for obtaining sufficient amounts of sleep.

**Effects of chronic sleep restriction**

Even though brief sleep periods can limit the cognitive deficits from cumulative sleep loss in the short-term (see Fig. 2), they fail to preserve optimal cognitive functioning in the long-term. This has been demonstrated in recent experiments of chronic sleep restriction [13,27,28]. In one of these studies [13], subjects were randomized to 14 days of restriction to 4-, 6-, or 8-hours time in bed per day. Fig. 3 shows results of this study for psychomotor vigilance lapses as measured with the PVT, averaged within days. Because of this averaging, changes within days resulting from the interaction of the homeostatic process and the circadian process are not visible in the figure, but more long-term changes in cognitive performance are clearly exposed. Compared with the control condition of 8 hours time in bed per day (in which subjects actually obtained approximately 7 hours of physiologic sleep), the sleep restriction conditions of 4 and 6 hours time in bed per day (dotted and thin curves, respectively) displayed progressive increases in psychomotor vigilance impairment. After 14 days of sleep restriction, the magnitude of impairment in the condition with 4 hours time in bed actually approached the daytime level of impairment observed during 88 hours of total sleep deprivation (see Fig. 1A).

These findings cannot be understood solely in terms of homeostatic pressure for sleep [13,29]. Based on the exponential nature of the homeostatic process, a new equilibrium would have been predicted to set in within a few days, when the dissipation of homeostatic pressure during restricted sleep should have become so much swifter (ie, exponentially faster) that it could compensate for the additional increase of homeostatic pressure during extended wakefulness each day. The results of chronic sleep restriction studies [13,27] have not
supported this prediction, as performance impairment continued to accumulate over days of sleep restriction. Two alternative models have been proposed: one in which the effects of sleep deprivation are described in terms of cumulative time of wake extension instead of a sleep homeostatic process [13], and one in which long-term changes in sensitivity to sleep loss are postulated [27,30]. New experiments are needed to determine which of these models best reflects the true nature of cognitive impairment from chronic sleep restriction.

It has been pointed out that the recuperation of performance capability appears to be slower after chronic sleep restriction (Fig. 3) than after acute total sleep deprivation (see Fig. 1) [27]. As can be seen in Fig. 3 (right side), two nights with 8 hours time in bed for recovery sleep appeared to only partially reduce the psychomotor vigilance lapses from the prior 14 days of sleep restriction (6 hours or 4 hours time in bed per day). The subjects in the condition with...
6 hours time in bed (thin curve) seemed to recuperate less than those in the condition with 4 hours time in bed (dotted curve), although the difference between conditions was not significant on the first recovery day ($F_{1,24} = 0.10$, $P = 0.76$) or on the second recovery day ($F_{1,24} = 0.32; P = 0.58$). However, the data shown are averages over subjects, so it is possible that only a few subjects with heightened vulnerability to sleep loss (as discussed later) created the appearance of incomplete recovery. Available data sets [13,27] have not resolved this issue definitively, and further studies are underway.

The performance-impairing effects of chronic sleep restriction can also be seen in a variety of cognitive functions other than psychomotor vigilance, but many of the performance tasks used to measure these other cognitive functions exhibit practice effects. An example is given in Fig. 4, which shows performance on a serial addition/subtraction task (SAST) [31] in the same study as depicted in Fig. 3 [13]. The extent of the practice effect is exposed in Fig. 4 by the considerable performance improvement over days (ie, upward trend) for the control condition (8 hours time in bed per day). The performance improvement was moderated in the conditions with less than 8 hours time in bed per day (Fig. 4, thin and dotted curves) because of the effect of cumulative sleep loss, but performance on the SAST did not decrease over days even in the condition with 4 hours time in bed per day. Thus, if performance changes resulting from the practice effect had been overlooked, and a control condition (8 hours time in bed per day) had not been included in the study, a false conclusion could have

![Fig. 4](image_url)

Fig. 4. Data from the same experiment as shown in Fig. 3, but for a different measure of cognitive function. The number of correct responses on a serial addition/subtraction task is displayed (as daily means), expressed relative to baseline performance on day 3 (ie, on the first condition day, before the first restricted sleep period). In this figure, upwards on the ordinate corresponds to performance improvement (not impairment as in the other figures). In the control condition with 8 hours time in bed per day, performance improved steadily over days because of the practice effect associated with the serial addition/subtraction task. In the condition with 6 hours time in bed, the improvement over days was attenuated because of the cumulative sleep loss, and in the condition with 4 hours time in bed, almost no improvement was observed—until after the first recovery sleep. Even after two recovery sleep periods at the end of the study, both sleep restriction conditions exhibited reduced task performance relative to the control condition.
been drawn from these data, suggesting that chronic sleep restriction did not adversely affect cognitive functioning, although in reality it did. It has not yet been established whether chronic sleep restriction merely affects performance output; and/or for cognitive tasks with a practice effect, whether chronic sleep loss interferes with the actual learning of the task. It could be argued that if the data in Fig. 4 simply reflected a reduction in performance output that masked the underlying practice effect (which by itself continued unaltered regardless of sleep loss), then at the end of the study, after recovery sleep, cognitive performance levels in the two sleep restriction conditions should have approached those in the control condition. No evidence of this is displayed in Fig. 4; the difference among the three conditions in SAST performance (expressed relative to each subject’s baseline performance) was significant on the first recovery day ($F_{2,31} = 5.35, P = 0.010$) and the second recovery day ($F_{2,31} = 3.70, P = 0.036$). This suggests that chronic sleep loss may have interfered with the practice effect proper.

Such an adverse effect of sleep loss on the cognitive benefit of practice would be in line with recent discoveries that sleep deprivation may reduce the brain’s ability to learn performance tasks [32]. Yet, some further evidence has indicated that even brief sleep periods could suffice for learning [33]. This matter may be dependent on the structure of sleep and on the nature of the performance task [34]. Psychomotor vigilance performance as measured with the PVT does not show any significant practice effect [13], and therefore task learning was not a notable factor for the psychomotor vigilance results shown in Figs. 1–3.

### Individual differences in vulnerability to sleep loss

Humans have been found to differ substantially in the degree of cognitive performance impairment they suffer from sleep loss [35–37], whether under conditions of acute total sleep deprivation or chronic partial sleep deprivation [13]. This is illustrated in Fig. 5, which shows PVT lapse data from a study in which subjects repeatedly underwent 36 hours of continuous wakefulness, that is, on two separate occasions [37]. Let us consider the group-average performance profiles in this study first.

The two thick, solid curves in Fig. 5 represent the group-average changes in psychomotor vigilance lapses during the two exposures to sleep deprivation. As expected, the shape of these curves resembles the first part of the 88-hour sleep deprivation curve shown in Fig. 1A (up to approximately hour 70 on the abscissa). Nevertheless, the average number of lapses after any given duration of wakefulness was much greater in the repeated 36-hour sleep deprivation study (see Fig. 5) than in the 88-hour sleep deprivation study (see Fig. 1A), because of the difference in task duration between the studies (20 minutes versus 10 minutes, respectively). Lapses on the PVT increase progressively with time on task [15], so that the number of lapses that may occur in a 20-minute PVT bout is much greater than twice the number occurring in a 10-minute PVT.
bout (under otherwise identical circumstances). Regardless of this difference in the absolute number of performance lapses, however, the normal interaction of the homeostatic process with the circadian process was observed in the profile of performance changes during the repeated 36-hour sleep deprivation study, and this profile was very similar between the two sleep deprivations (see Fig. 5, solid curves).

Fig. 5 also shows the profiles of performance changes in two individual subjects participating in the repeated sleep deprivation study. One subject, indicated with the dotted curves, was relatively resistant to the 36 hours of sleep deprivation, showing only a small dip in performance in the early morning hours (ie, when the circadian pressure for wakefulness should have been low). This same pattern was observed during both sleep deprivations, indicating that it was not a chance observation but a characteristic of the individual at hand. Another subject, indicated with the thin curves, was relatively vulnerable to the effects of sleep deprivation. This was seen at all times of wakefulness past the normal waking day, and for both exposures to sleep deprivation. This same phenomenon was observed across all subjects involved in the study; compared with each other, they varied substantially in the magnitude of performance.
impairment during sleep deprivation, but the performance profile was highly replicable within each individual. In fact, as much as 67.5% of the variance in the psychomotor vigilance data was explained by consistent individual differences [37].

Each 36-hour sleep deprivation session was preceded by seven consecutive days with time in bed extended to 12 hours per night. Thus, the differences among individuals observed during sleep deprivation could not have been caused by uncontrolled differences in prior amounts of sleep (or sleep insufficiency) [37]. The issue of prior sleep amounts (ie, “sleep history”) was investigated further by having each of the subjects undergo a third 36-hour sleep deprivation, which was preceded by seven consecutive days with time in bed restricted to 6 hours per night. As implied by the data in Fig. 3 (days 4–10), this chronic sleep restriction would have been expected to induce marked susceptibility to performance impairment even before the 36-hour sleep deprivation began. It turned out, however, that the effect of the seven prior days of sleep restriction (at 6 hours time in bed per day) on psychomotor vigilance impairment during total sleep deprivation was small compared with the considerable individual differences consistently observed during the other two sleep deprivation sessions (and again noticed during the third). This finding indicates that individual differences in psychomotor vigilance impairment from sleep loss are a robust trait, which has been dubbed differential vulnerability [37].

The origin of the trait individual differences in performance impairment from sleep loss has remained unclear. The individual subjects’ data in Fig. 5 would suggest that differences in psychomotor vigilance performance at baseline (ie, during the first 12 hours of continuous wakefulness) might predict the performance response to sleep deprivation (ie, during the last 24 hours of continuous wakefulness). However, less than 25% of the between-subjects variance in PVT performance during sleep deprivation was actually explained by individual differences at baseline; the correlation between baseline performance (the average over the first 12 hours of wakefulness) and the response to sleep deprivation (the average over the last 24 hours of wakefulness) for subjects’ first exposure to sleep deprivation (preceded by 7 days of sleep extension to 12 hours time in bed) was $r = 0.486 \ (P = 0.025)$. A search for better predictors of differential vulnerability to sleep loss is ongoing.

Individual differences in performance impairment during sleep deprivation were also noticed in other cognitive functions such as working memory, but when individuals were ranked by their degree of vulnerability, their rank order was found to be different for psychomotor vigilance performance than for other performance measures investigated thus far. It appears, therefore, that psychomotor vigilance is a distinct aspect of cognitive functioning, possibly subserved by specific neurocognitive pathways in the brain. These pathways appear to clearly reflect the interaction of sleep homeostatic and circadian neurobiology. As such, they are of relevance to more complex tasks [16], such as motor vehicle operation and athletic performance.
References


