Systematic Interindividual Differences in Neurobehavioral Impairment from Sleep Loss: Evidence of Trait-Like Differential Vulnerability

Hans P. A. Van Dongen, PhD; Maurice D. Baynard, MS; Greg Maislin, MS, MA; David F. Dinges, PhD

Unit for Experimental Psychiatry, Division of Sleep and Chronobiology, Department of Psychiatry, and Center for Sleep and Respiratory Neurobiology, University of Pennsylvania School of Medicine, Philadelphia

Objectives: To investigate interindividual differences in neurobehavioral deficits during sleep deprivation, and to establish to what extent the neurobehavioral responses to sleep loss are a function of sleep history versus trait-like differential vulnerability.

Design: Individuals were exposed to sleep deprivation on 3 separate occasions in order to determine the stability of interindividual differences in neurobehavioral impairment.

Setting: The sleep-deprivation experiments were conducted under standardized laboratory conditions with continuous monitoring of wakefulness. Each subject underwent a laboratory-adaptation session before entering the sleep-deprivation phase of the study.

Participants: A total of 21 healthy adults (aged 21–38 years) completed the experiment.

Interventions: Subjects came to the laboratory 3 times at intervals of at least 2 weeks. During each laboratory session, they underwent neurobehavioral testing every 2 hours during 36 hours of total sleep deprivation, which was preceded by baseline sleep and followed by recovery sleep. In the week prior to each sleep-deprivation session and on the baseline night in the laboratory, subjects were required to either restrict their sleep to 6 hours per day (prior sleep restriction condition) or to extend their time in bed to 12 hours per day (prior sleep extension condition), so as to experimentally manipulate sleep history (in randomized counterbalanced order).

Results: There was strong evidence that interindividual differences in neurobehavioral deficits during sleep deprivation were systematic and trait-like. The magnitude of interindividual variability was substantial relative to the magnitude of the effect of prior sleep restriction (which on average involved a reduction of 4.1 hours sleep per day, compared to prior sleep extension, for 7 days). Overall, interindividual differences were not explained by subjects' baseline functioning or a variety of other potential predictors. Interindividual variability clustered on 3 distinct neurobehavioral dimensions: self-evaluation of sleepiness, fatigue, and mood; cognitive processing capability; and behavioral alertness as measured by sustained attention performance.

Conclusions: Neurobehavioral deficits from sleep loss varied significantly among individuals and were stable within individuals. Interindividual differences in neurobehavioral responses to sleep deprivation were not merely a consequence of variations in sleep history. Rather, they involved trait-like differential vulnerability to impairment from sleep loss, for which neurobiologic correlates have yet to be discovered.

Key Words: individual differences, differential vulnerability, neurobehavioral functions, cognitive performance, sleepiness, sleep deprivation, sleep history, trait variability, healthy adults

Citation: Van Dongen HPA; Baynard MD; Maislin G; Dinges DF. Systematic interindividual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability. *SLEEP* 2004; 27(3):423-33.

INTRODUCTION

THE EFFECTS OF SLEEP LOSS ON WAKING NEUROBEHAVIORAL FUNCTIONS HAVE BEEN A CRITICAL TOPIC OF INVESTIGATION IN BOTH BASIC AND CLINICAL SLEEP RESEARCH. In humans, sleep loss produces a range of fundamental neurocognitive deficits such as reductions in vigilance, working memory, and executive function.[1,2] Large interindividual differences in these deficits have been observed, however, accounting for a substantial portion of the variance.[3] Yet, interindividual variability in responses to sleep deprivation has been mostly overlooked in the scientific literature. Because recent sleep history can have a significant effect on neurobehavioral functions,[4,5] it could be hypothesized that interindividual differences in sleep history may explain the observed differences in responses to subsequent sleep deprivation. Still, variability in sleep history does not explain the persistent interindividual differences encountered in laboratory-based sleep-deprivation studies that control for prior sleep.[4,6] To date, only 4 studies[6–9] have systematically evaluated interindividual variability in neurobehavioral deficits from sleep loss by repeatedly subjecting individuals to sleep deprivation, and none of these studies have properly quantified the magnitude and nature of the observed variability.

Wilkinson[7] and Webb and Levy[8] were the first to study multiple exposures to total sleep deprivation—in 12 and 6 subjects, respectively—and to observe anecdotally that there were large, consistent interindividual differences in performance deficits resulting from sleep loss. These studies did not provide conclusive evidence of predictable interindividual differences, however, because the laboratory environments were not fully controlled, circadian times of measurement were not aligned, and subjects' activity-rest schedules prior to sleep deprivation were not standardized. Furthermore, neither Wilkinson nor Webb and Levy reported quantitative results of the interindividual differences in performance deficits from sleep loss.

Leproult and colleagues[6] addressed the issue of interindividual variability more quantitatively in an experiment involving sleep deprivation of 8 subjects twice under similar constant routine conditions. Subjects were tested hourly on a selective attention task, a sustained attention task, and a visual analog scale for global vigor. Parametric and nonparametric correlations between the magnitudes of impairment, calculated as the difference between the maximum and the minimum in the temporal profile, in the first versus the second exposure to sleep deprivation were used to estimate the stability of interindividual differences in the response to sleep deprivation. For global vigor, the correlation coefficients were reported to be 0.95 (parametric) and 0.90 (nonparametric);

Disclosure Statement
No significant financial interest/other relationship to disclose.

Submitted for publication December 2003
Accepted for publication February 2004
Address correspondence to: Hans P. A. Van Dongen, PhD, Division of Sleep and Chronobiology, Department of Psychiatry, University of Pennsylvania School of Medicine, 1019 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021; Tel: (215) 573-5866; Fax: (215) 573-6410;
E-mail: vdongen@mail.med.upenn.edu

for reaction time on the combined attention tasks, the correlation coefficients were 0.93 (parametric) and 0.88 (nonparametric). Various methodologic issues make it difficult to interpret these findings as true quantitative evidence of the reproducibility of interindividual differences in neurobehavioral impairment from sleep loss. The data were smoothed before analysis, which reduced the variance *within* subjects relative to the variance *between* subjects and, therefore, artificially increased the apparent interindividual variability. The data were then normalized by expressing them relative to the overall minimum or maximum (ie, normalization by division), which resulted in disproportional propagation of measurement error into the observed effects and the estimates of interindividual variability. Finally, the use of the difference between maximum and minimum as a measure of the magnitude of impairment, with minimum and maximum assessed separately in each of the 2 constant routines, further confounded the results. With this approach, the investigators introduced time as an additional degree of freedom in the analyses of interindividual variability, without properly accounting for it in statistical results.

Correlation statistics do not readily generalize to account for multiple sources of variance. A more flexible, statistically valid approach to quantifying interindividual variability in neurobehavioral responses to sleep loss is the intraclass correlation coefficient (ICC).[10] This metric is based on variance components analysis,[11] involving the explicit separation of *within-subjects* variance and *between-subjects* variance in data derived from repeated exposures to an experimental intervention. The ICC expresses the proportion of variance in these data that is explained by systematic interindividual variability. We previously used the ICC in a study of 10 subjects exposed twice to 40 hours of sleep deprivation under laboratory control.[9,10] Subjects were tested every 2 hours on a psychomotor vigilance task (PVT), and performance decrements were computed as the number of performance lapses in the last 10 hours of sleep deprivation (ie, response to sleep loss) minus the number of lapses at the same clock times 24 hours earlier (ie, baseline). The value of the ICC was 0.58, indicating that 58% of the variance in the effect of sleep loss could be attributed to trait-like interindividual variability. We may have underestimated the magnitude of stable interindividual differences in this experiment,[12] however, because physical activity (eg, ambulation) and environmental stimulation (eg, social interaction) were available to the subjects in 1 of the 2 exposures to sleep deprivation and disallowed in the other. Thus, the within-subjects variance included systematic differences between the 2 sleep deprivations, thereby artificially reducing the value of the ICC.

In order to reliably investigate the stability of individual responses to sleep loss, state-dependent variance must be minimized among subjects as well as across exposures to sleep deprivation by standardizing demand characteristics, controlling environmental factors, and satiating any preexisting sleep debt, in an experiment involving repeated exposures to sleep deprivation.[10] Minimally, a strictly controlled laboratory study involving at least 2 exposures to total sleep deprivation (TSD) under carefully standardized circumstances can serve to assess the true magnitude of trait interindividual variability in vulnerability to neurobehavioral impairment from sleep loss. The importance of the interindividual differences in such an experiment may be difficult to interpret, however, unless the experiment also includes a condition involving manipulation of a conceptually meaningful state variable. This way, the magnitude of trait interindividual variability in responses to sleep loss can be evaluated relative to the magnitude of change in the responses to sleep loss caused by the manipulation of state. In the present study, we selected sleep history to be the state variable in question, in order to investigate whether interindividual differences in neurobehavioral deficits during TSD can be attributed to variations in sleep history.

To manipulate sleep history, we required subjects to restrict their time in bed to 6 hours per day in the week before 1 exposure to laboratory sleep deprivation (PSR: *prior sleep restriction* condition), and to extend their time in bed to 12 hours per day in the week before 2 additional exposures to laboratory sleep deprivation (PSE: *prior sleep extension* conditions). We have shown previously that sleep restriction at 6 hours time in bed per day for 7 days leads to significant neurobehavioral deficits,[4] against which the magnitude of interindividual variability may be judged. In the present experiment, we quantified trait interindividual variability using objective and subjective neurobehavioral outcomes obtained from the 2 laboratory sleep deprivations preceded by sleep extension (by which we aimed to satiate any preexisting sleep debt) and compared this trait variability to the effect of PSR on neurobehavioral functions in the third laboratory sleep deprivation. The conditions occurred in randomized counterbalanced order. Subjects also underwent a laboratory-adaptation session before the exposures to sleep deprivation began.

We report the results of the first experimental study systematically quantifying the relative contributions of trait interindividual variability versus sleep history to the magnitude of neurobehavioral changes induced by sleep deprivation. Preliminary analyses of a small subset of the present data were briefly described in a recent review paper.[3]

METHODS

Subjects participated in *3 replications* of 36-hour TSD in a laboratory. The amount of sleep allowed at home during the week prior to each laboratory session was varied in randomized counterbalanced order between a PSE condition (ie, 12 hours time in bed per day), which occurred twice, and a PSR condition (ie, 6 hours time in bed per day), which occurred once.

Subjects

A total of 21 subjects completed the study. They were 12 men and 9 women, ranging in age from 21 to 38 years (mean ± SD: 29.5 ± 5.3 years). Since estimates of interindividual variability depend on the sample studied, it is important to characterize the population from which this sample was drawn. In order to be eligible for participation in the study, subjects met the following criteria: age from 21 to 40 years; physically and psychologically healthy, as assessed by physical examination and history; no clinically significant abnormalities in blood chemistry and urine samples and free of traces of drugs; good habitual sleep, between 6.5 and 8.5 hours in duration daily; regular bedtimes, getting up between 6:30 and 8:30 AM; neither extreme morning nor extreme evening type, as assessed by questionnaire[13]; no sleep or circadian disorder, as assessed by questionnaire[14] and polysomnography; no history of psychiatric illness and no previous adverse neuropsychiatric reaction to sleep deprivation; no history of alcohol or drug abuse; and no current medical or drug treatment (excluding oral contraceptives).

The study was approved by the Institutional Review Board of the University of Pennsylvania, and all subjects gave written informed consent.

Experimental Design

The experiment was conducted in a controlled laboratory environment in the General Clinical Research Center of the Hospital of the University of Pennsylvania. Before entering the experimental phase of the study, subjects underwent a laboratory-adaptation session, during which they practiced the neurobehavioral test battery and acclimated to sleeping in the laboratory while being monitored polysomnographically (12 hours time in bed). The adaptation session resembled the first part of the subsequent 3 laboratory TSD sessions.

Two weeks after the adaptation session, subjects returned to the laboratory for their first TSD session. They were scheduled to enter the laboratory at 3:00 PM, after which they were prepared for polysomnographic recording of subsequent sleep. They also practiced the neurobehavioral test battery once more. In the PSE condition, subjects went to bed at 10:00 PM for a 12-hour baseline sleep period that was recorded polysomnographically. In the PSR condition, however, they stayed awake 6 hours longer, performing the neurobehavioral tests at 10:00 PM,

midnight, and 2:00 AM; they then went to bed at 4:00 AM for a 6-hour sleep period that was also recorded polysomnographically. All sleep periods ended at 10:00 AM.

Following the baseline sleep period, subjects underwent TSD for 36 hours. Wakefulness was monitored continuously by trained staff members. Starting at 10:00 AM in the beginning of the TSD period, neurobehavioral tests were administered every 2 hours. At 10:00 PM at the end of the 36-hour TSD, subjects went to bed for a 12-hour recovery sleep opportunity that was recorded polysomnographically. At 10:00 AM the next morning, they performed the neurobehavioral tests a final time, took a shower (which was not allowed at any other times in the laboratory), and departed at around noon.

The conditions in the laboratory were strictly controlled in terms of neurobehavioral test schedules and activities. Subjects were allowed only nonvigorous activities between test bouts, and they had no interactions with people outside the laboratory. Light exposure was less than 50 lux during scheduled wakefulness and less than 1 lux during scheduled sleep. Ambient temperature was maintained at 21°C ± 1°C. Standardized meals were given at 7:00 PM (and 11:00 PM and 3:00 AM in the PSR condition) before baseline sleep; at 11:00 AM, 3:00 PM, 7:00 PM, 11:00 PM, 3:00 AM, 7:00 AM, 11:00 AM, 3:00 PM, and 7:00 PM during the TSD period; and at 11:00 AM after recovery sleep. Food was carefully controlled in terms of calories and nutrients (proteins, fats, and carbohydrates). The amount of food subjects received during the 36 hours of TSD matched their normal 2-day caloric requirement based on height and weight. The laboratory TSD session was repeated after an interval of at least 2 weeks and was conducted a third time after another interval of at least 2 weeks. Subjects were not allowed to use any caffeine, alcohol, tobacco, or medications during the week before each of the TSD sessions—as verified by means of urine screens—and during the laboratory stays.

During the week prior to each laboratory TSD, subjects were required to adjust their daily bedtimes at home in accordance with the assigned experimental condition. In the PSE condition, subjects were required to stay in bed from 10:00 PM until 10:00 AM each day (ie, a 12-hour daily sleep opportunity). In the PSR condition, subjects' sleep opportunities were restricted to the period from 4:00 AM until 10:00 AM each day (ie, a 6-hour daily sleep opportunity). Subjects were instructed not to engage in any safety-sensitive tasks (such as driving) while in this condition. Daytime napping was not allowed in either the PSE condition or the PSR condition. The baseline sleep period in the laboratory matched the scheduled sleep times at home (ie, either 6 hours or 12 hours duration), in accordance with the assigned experimental condition. Only 1 of the 3 laboratory TSD sessions was preceded by the PSR condition, while the other 2 TSD sessions were preceded by the PSE condition, in randomized counterbalanced order. Thus, one third of the subjects encountered the PSR condition before their first sleep-deprivation session, one third encountered it before their second sleep deprivation, and one third encountered it last.

Neurobehavioral Assessments

Subjects underwent computerized neurobehavioral tests every 2 hours during scheduled wakefulness. The neurobehavioral test battery contained the following objective and subjective evaluations: *Karolinska Sleepiness Scale (KSS)*, a Likert-type rating scale of subjective sleepiness[15]; computerized *visual analog scales of fatigue* (*VAS-F*) *and mood (VAS-M)*, anchored by "fresh" and "exhausted" and by "elated" and "depressed", respectively[16]; *serial addition/subtraction task (SAST)*, a cognitive task taken from the Walter Reed performance assessment battery[17]; *digit symbol substitution task (DSST)*, a computerized version of the cognitive performance task of the same name in the Wechsler Adult Intelligence Scale[18]; *critical tracking task*, a manual tracking task[19] on which we do not report here because of equipment problems (which did not interfere with the presentation of the task); *word detection task (WDT)*, a shortened version of a signal detection task used in the earlier repeated sleep-deprivation study by Webb and Levy[8]; *repeated acquisition of response sequences task (RARST)*, a cognitive task used previously in research on hypnotics[20]; *psychomotor vigilance task (PVT)*, a cognitive test of sustained attention that uses reaction times to measure behavioral alertness[21]; and *performance and effort rating scales (PERF and EFF)*, a pair of computerized scales on which subjects rated their performance and the effort required of them to keep performing.[22]

Table 1 lists the neurobehavioral evaluations, as well as the acronyms by which we refer to them. The KSS and the VAS-F and VAS-M were included twice (ie, at the start of the test battery, and near the end). Table 1 also shows the outcome measures yielded by the different tests and the time it took to complete each test. The duration of the tests was controlled by the test computer; it took 1 hour to finish the entire battery. Subjects used the same desktop test computer during the entire study. They were seated throughout all neurobehavioral testing periods and were behaviorally monitored continuously. Subjects were instructed to perform to their best level and to use compensatory effort to keep up performance. One subject gave identical responses to the subjective neurobehavioral evaluations each time. Because we could not be certain that these responses were truthful, and because they might inflate estimates of trait variability, we removed this subjects' self-report outcomes from the data set. Two other subjects showed instances of performance noncompliance (ie, excessive numbers of false starts) on the PVT. These subjects' PVT outcomes were also removed from the data set.

For the purposes of statistical analysis, the neurobehavioral data were averaged over the last 24 hours of each of the 36-hour TSD periods. Thus, for each TSD period, a single value summarizing the response to sleep deprivation was computed, averaged over the circadian cycle. As a result, there were 3 summary response values per subject for every neurobehavioral outcome variable. Two of these were associated with the PSE conditions, and 1 was associated with the PSR condition. The 13 neurobehavioral outcome variables were also averaged across the first 12 hours of each of the TSD periods in order to compute the corresponding baseline daytime levels of neurobehavioral functioning for each subject.

Table 1—Neurobehavioral evaluations in the test battery

Neurobehavioral Test	Acronym	Outcome Measure	Completion Time
Karolinska Sleepiness Scale (1st administration)	KSS-1	Subjective sleepiness score (1–9)	1 min
Visual analog scale of fatigue (1st administration)	VAS-F1	Subjective fatigue score (1–9)	18 sec
Visual analog scale of mood (1st administration)	VAS-M1	Subjective mood score (1–9)	18 sec
Serial addition/subtraction task	SAST	Number of correct responses	6.5 min
Digit symbol substitution task	DSST	Number of correct responses	6.5 min
Critical tracking task	CTT	No outcome measure available	7 min
Word detection task	WDT	Number of correct responses	8 min
Repeated acquisition of response sequences task	RARST	Number of acquisitions	8 min
Psychomotor vigilance task	PVT	Number of lapses (RT ≥ 500 ms)	20 min
Karolinska sleepiness scale (2nd administration)	KSS-2	Subjective sleepiness score (1–9)	1 min
Visual analog scale of fatigue (2nd administration)	VAS-F2	Subjective fatigue score (1–9)	18 sec
Visual analog scale of mood (2nd administration)	VAS-M2	Subjective mood score (1–9)	18 sec
Performance rating scale	PERF	Performance evaluation (1-7)	24 sec
Effort rating scale	EFF	Expended effort evaluation (1–4)	24 sec

From left to right, the table lists the different tests, the acronyms we used for them, their outcome measures, and the time needed to complete them. It took 1 hour to finish the whole test battery.

Other Measurements

The adaptation night and the baseline and recovery nights of the 3 laboratory TSD sessions were all recorded polysomnographically (Vitaport 3; TEMEC Instruments BV, Kerkrade, The Netherlands). Subjects were given the same digital recorder each time they visited the laboratory. The polysomnographic montage included frontal (Fz), central (C3, C4), and occipital (Oz) electroencephalogram (referenced against A1/A2), bilateral electrooculogram, and submental electromyogram; we also recorded the electrocardiogram. All sleep recordings were scored using conventional criteria,[23] blind to conditions (only the 6-hour baseline recordings in the PSR condition could be recognized) and blind to subjects. Out of the 63 baseline recordings (ie, 3 baseline records for each of the 21 subjects), 10 could not be scored reliably because of equipment problems (3 of these were 6 hours in duration as per the PSR condition).

Throughout the laboratory sessions, subjects' core body temperature was recorded by means of a rectal probe (Series 400; Yellow Springs Instrument, Yellow Springs, OH) connected to a digital recorder (Mini-Logger; Mini Mitter Co., Inc., Bend, OR), at the rate of 10 samples per hour. A 2-harmonic sinusoidal regression model[24,25] was fitted to the temperature data acquired during the 36-hour TSD periods. Half the difference between the maximum and the minimum in the fitted regression model was used as an estimate of circadian amplitude. The timing of the minimum in the model was used as a marker of circadian phase. Data loss due to uncontrollable variability in probe position and as a result of equipment problems was substantial. Since there is considerable stability in circadian phase[26] and amplitude[a] over long intervals, a single complete temperature curve per subject was used for the purpose of the present investigation.

Subjects' activity was recorded during the entire experiment, both during laboratory stays and in the week before each laboratory session (ie, during the assigned PSR and PSE conditions at home). The wrist actigraph used for this purpose (Actiwatch-L; Mini Mitter Co., Inc.), which was worn on the nondominant arm, also recorded light exposure. Subjects received the same actigraph each time they were in one of the assigned sleep conditions and each time they were in the laboratory. Automated analyses of the activity data (Actiware Sleep 3.1; Mini Mitter Co., Inc.) yielded initial estimates of daily sleep durations during the assigned sleep conditions at home. During each week of PSR or PSE, subjects also kept a diary to record their bedtimes, and they called a time-stamped telephone recorder every morning and every night to report their bedtimes. Adherence to the PSR and PSE conditions could thus be verified.

Prior to the experiment, subjects filled out a number of questionnaires providing information on various demographics, sleep–wake-related variables, and psychosocial traits, which the scientific literature suggested might predict individuals' responses to sleep deprivation. In addition to surveys of sleep habits, medical background, and demographics, these questionnaires included the Composite Scale of Morningness/Eveningness,[13] the Epworth Sleepiness Scale,[27] the Pittsburgh Sleep Quality Index,[28] the Sleep Disorders Questionnaire,[14] the Conflict-Stress Questionnaire,[29] the Eysenck Personality Inventory,[30] the Marlowe-Crowne Social Desirability Scale,[31] the Penn State Worry Questionnaire,[32] the Spielberger State-Trait Anxiety Inventory,[33] the Survey of Work Styles,[34] and the Myers-Briggs Type Indicator.[35]

Statistical Analyses of Sleep Duration

Analyses focused first on assessing whether the experimental manipulation of sleep history was successful. Daily sleep durations in the 7 days before each of the 3 laboratory TSD sessions were estimated from the bedtimes that subjects reported by telephone, from the diaries that they kept, and from the automated analyses of their actigraphic records. For the last night before TSD, which subjects spent in the laboratory, these estimates were replaced by the total sleep times determined from visual scoring of the polysomnogram. The estimated sleep durations were then averaged over the 7 days prior to each of the TSD sessions.

To verify the experimental control over sleep history, we compared average daily sleep durations between the 2 PSE conditions using a linear mixed-model analysis of variance.[36,37] Furthermore, the average sleep durations in the PSR condition and the 2 PSE conditions were compared by means of a linear mixed-model regression analysis[36,37] involving—in addition to an intercept—a fixed effect for sleep restriction (relative to sleep extension), a fixed effect for the TSD session number (expressing whether data were taken from the week before the first, second, or third exposure to TSD), an interaction of session number by sleep restriction, and a random effect for the intercept. Because no significant effects were found for experimental session number and interaction of session number by sleep restriction (ie, there were no order effects), these terms were dropped. The reduced mixed model was used to estimate the difference in average daily sleep duration between the PSR condition and the PSE conditions. To verify the robustness of the findings, these analyses were repeated using only the total sleep time data for the laboratory baseline nights.

Statistical Analyses of Trait-Like Differential Vulnerability

For the analyses of trait-like differential vulnerability to neurobehavioral impairment during TSD, we considered only the responses to sleep deprivation (averaged over the last 24 hours) in the 2 PSE conditions. We separated the between-subjects variance σ^2_{bs} from the within-subjects variance σ^2_{ws} in these data by means of variance components analyses,[11] which we implemented as linear mixed-model analyses of variance with fixed-effects corrections for order effects (ie, for the placement of the PSE conditions relative to the PSR condition in each subject). All mixed-model analyses in this study were performed with the restricted maximum likelihood method (SAS 8.02; SAS Institute Inc., Cary, NC). To quantify trait-like interindividual variability, we assessed the ICC, which was computed as the ratio of the between-subjects variance σ^2_{bs} to the total variance $\sigma^2_{bs} + \sigma^2_{ws}$ after removing the estimated fixed order effects.

The ICC values were interpreted using published benchmark ranges,[38] which can be interpreted as corresponding to increasing stability of observed interindividual differences: "slight" (0.0–0.2); "fair" (0.2–0.4); "moderate" (0.4–0.6); "substantial" (0.6–0.8); and "almost perfect" (0.8–1.0). Statistical significance of the ICC values was assessed by means of a Wald Z test of the between-subjects variance. Order effects were tested for statistical significance using an omnibus F test (4,18 degrees of freedom). To determine if observed interindividual differences in responses to TSD were the result of interindividual differences at baseline, we repeated the mixed-model analyses of variance with baseline daytime neurobehavioral functioning as a covariate and recomputed the ICC values. The statistical significance of the baseline covariate was investigated with a t test (17 degrees of freedom).

Provided that the interindividual differences in the response to TSD were robust across the 2 PSE conditions (ie, ICC values in the "substantial" or "almost perfect" ranges, which means ICC > 0.6), we averaged the neurobehavioral responses over the 2 TSD sessions with PSE (ie, with satiation of any preexisting sleep debt). These averages were considered to investigate whether or not interindividual differences were idiosyncratic to each neurobehavioral test. A (principal) factor analysis with orthogonal varimax rotation (SAS 8.02; SAS Institute, Inc.) was performed on all 13 neurobehavioral outcomes. We inspected the scree plot of eigenvalues to determine how many factors should be retained before rotation in order to explain most of the variance in the data set. Given the comparatively small sample size of the study (n = 21), we interpreted only factor loadings greater than 0.5.

Statistical Analyses of Potential Correlates of Differential Vulnerability

Subjects' neurobehavioral responses to TSD averaged across the 2

[a]In a constant routine study that we conducted,[26] the circadian amplitude of core body temperature was found to be highly stable, with an intraclass correlation coefficient (ICC) of 0.907 (ie, near-perfect replicability).

PSE conditions were used to investigate whether interindividual differences were associated with demographic parameters or sleep–wake-related variables. We considered subjects' age, sex, body mass index, handedness, self-reported habitual sleep duration (both on weekdays and on weekends or days off), prestudy overall sleepiness (Epworth Sleepiness Scale[27]), pre-study sleep quality (Pittsburgh Sleep Quality Index[28]), circadian preference (Composite Scale of Morningness/Eveningness[13]), and phase and amplitude of the circadian rhythm in core body temperature. Stepwise linear regression was performed (SPSS 11.5.2.1; SPSS Inc., Chicago, IL) to evaluate all 11 demographic and sleep–wake-related variables as candidate predictors (inclusion criterion: $\alpha = .05$; removal criterion: $\alpha = .10$).

We also investigated whether the interindividual differences were a function of psychosocial traits in the subject sample. We considered questionnaire assessments for psychoticism, extraversion, neuroticism, lying (Eysenck Personality Inventory[30]); extraversion/introversion, sensing/intuition, thinking/feeling, judging/perceiving (difference scores on the Myers-Briggs Type Indicator[35]); social desirability (Marlowe-Crowne Social Desirability Scale[31]); worrying (Penn State Worry Questionnaire[32]); stressful conditions, stress symptoms, relaxation methods (Conflict-Stress Questionnaire[29]); trait anxiety (Spielberger State-Trait Anxiety Inventory[33]); and impatience, anger, work involvement, time urgency, job dissatisfaction, competitiveness, type A personality (Survey of Work Styles[34]). Stepwise linear regression was performed to evaluate all of the 21 subscale scores as candidate predictors.

Statistical Analyses of Sleep History Effects

In order to assess the effect of the PSR condition (relative to the PSE conditions) on impairment during the 36-hour sleep-deprivation periods, we considered each subject's neurobehavioral responses to sleep deprivation (averaged over the last 24 hours) for the PSR condition and the 2 PSE conditions simultaneously. Since subjects were not anticipated to sleep for the full 12 hours time in bed in the PSE conditions, we performed an "as treated" analysis using actual sleep durations averaged over the 7 days before each of the 3 TSD sessions. For every neurobehavioral outcome of TSD, we constructed a mixed regression model containing—in addition to an intercept—a fixed effect for the average sleep duration prior to each of the TSD sessions, a fixed effect for the TSD session number (ie, first, second, or third exposure to TSD), a fixed interaction of session number by sleep duration, a random effect for the intercept, and a random effect for the impact of average sleep duration prior to each of the TSD sessions. Because the latter did not reach statistical significance for any of the 13 neurobehavioral tests, it was removed from the model in order to enhance statistical power. Likewise, as the interaction of session number by sleep duration did not reach statistical significance, it was dropped from the regression model as well, which led to a reduced mixed regression model.

Using the reduced model, the magnitude β of the effect of sleep history was assessed as the change in neurobehavioral functioning during TSD per additional hour of daily sleep obtained on average in the prior 7 days. The statistical significance of β was evaluated with a t test (39 degrees of freedom). Order effects (independent of sleep history) were tested for statistical significance using an omnibus F test (2,39 degrees of freedom). The between-subjects variance σ^2_{bs} for the intercept, representing systematic interindividual differences in the response to TSD (independent of sleep history), was tested for statistical significance by means of a Wald Z test. To verify the robustness of the findings, these analyses were repeated using only the total sleep times of the laboratory baseline nights.

The population-average effect γ of the experimental manipulation of sleep history in this study was determined by multiplying the average magnitude of the manipulation (ie, the difference in average daily sleep duration for the 7 days before TSD between the 2 PSE conditions and the PSR condition) with the estimated value of β (ie, the change in neurobehavioral functioning during TSD per additional hour of daily sleep obtained on average in the prior 7 days). In addition, the magnitude of systematic interindividual variability was estimated by the population standard deviation of neurobehavioral impairment during TSD, that is, the square root σ_{bs} of the between-subjects variance σ^2_{bs} as assessed simultaneously with the sleep history effect. We compared γ with σ_{bs} to determine if the effect of sleep history was dominant over trait-like interindividual variability, or vice versa.

Finally, we considered the population distribution of neurobehavioral impairment during TSD in the PSE conditions and investigated how much this distribution was shifted due to the experimental manipulation of sleep history in the PSR condition. On the basis of the response to TSD on a given neurobehavioral test in the PSE conditions, subjects were tentatively categorized as "resistant," which we defined as less impaired during TSD than the average individual, or as "vulnerable," which we defined as more impaired during TSD than the average individual. We judged the effect of sleep history to be substantial as a determinant of neurobehavioral impairment from sleep loss if the PSR condition caused the population distribution to shift so much that at least 50% of the subjects originally classified as resistant no longer belonged to that category. Based on the statistical properties of an assumed normal distribution, this was derived to be the case if $|\gamma| / \sigma_{bs} \geq 0.68$.

RESULTS

Manipulation of Sleep History

The success of our experimental manipulation of sleep history can be evaluated by the average daily sleep durations in the 7 days prior to each TSD session. In the PSR condition, the mean ± SD of the average daily sleep duration was 4.6 ± 0.8 hours; in subjects' first exposure to the PSE condition, this was 8.5 ± 1.0 hours, and in the second exposure to the PSE condition, it was 8.8 ± 1.3 hours. Mixed-model analysis of variance showed no significant difference in the average daily sleep duration between the 2 PSE conditions ($t_{18} = 1.25$, $P = .23$). Mixed-model regression analysis of the average sleep duration in the PSR condition and the 2 PSE conditions showed no significant effect of experimental session number ($F_{2,37} = 1.51$, $P = .24$) and no significant interaction of session number by condition ($F_{2,37} = 2.35$, $P = .11$), indicating that there were no significant order effects on the average sleep durations in this study. The 2 nonsignificant terms were dropped, and the mixed-model regression analysis was performed again. The result confirmed that subjects slept significantly less ($t_{41} = –17.49$, $P < .001$) in the PSR condition than in the PSE conditions; the difference (mean ± SEM) was 4.1 ± 0.2 hours.

Focusing just on the laboratory baseline nights before each of the 36-hour TSD periods, the mean ± SD of total sleep time was observed to be 5.3 ± 0.4 hours in the PSR condition, 9.2 ± 1.8 hours in the first exposure to the PSE condition, and 9.3 ± 1.4 hours in the second exposure to the PSE condition. There was no significant difference in total sleep time between the 2 PSE conditions ($t_{12} = 0.21$, $P = .83$). There were also no significant order effects among the PSR and PSE conditions (effect of experimental session number: $F_{2,28} = 1.34$, $P = .28$; interaction of session number by sleep restriction: $F_{2,28} = 0.21$, $P = .81$). Subjects slept 3.9 ± 0.3 hours (mean ± SEM) less in the baseline night of the PSR condition than in the baseline nights of the PSE conditions (significantly different: $t_{32} = –12.38$, $P < .001$). These results show that the experimental manipulation of sleep history was successful, and similar whether evaluated on the basis of sleep duration in the laboratory baseline nights or on the basis of average daily sleep duration in the 7 days before TSD.

Trait-Like Differential Vulnerability

To investigate differential vulnerability to neurobehavioral impairment during 36 hours of TSD, we used only data from the 2 PSE conditions (ie, 12 hours time in bed in the week prior to TSD). Figure 1 illustrates these data for 3 different neurobehavioral outcomes. Controlling for order effects, we separated the between-subjects variance σ^2_{bs} from

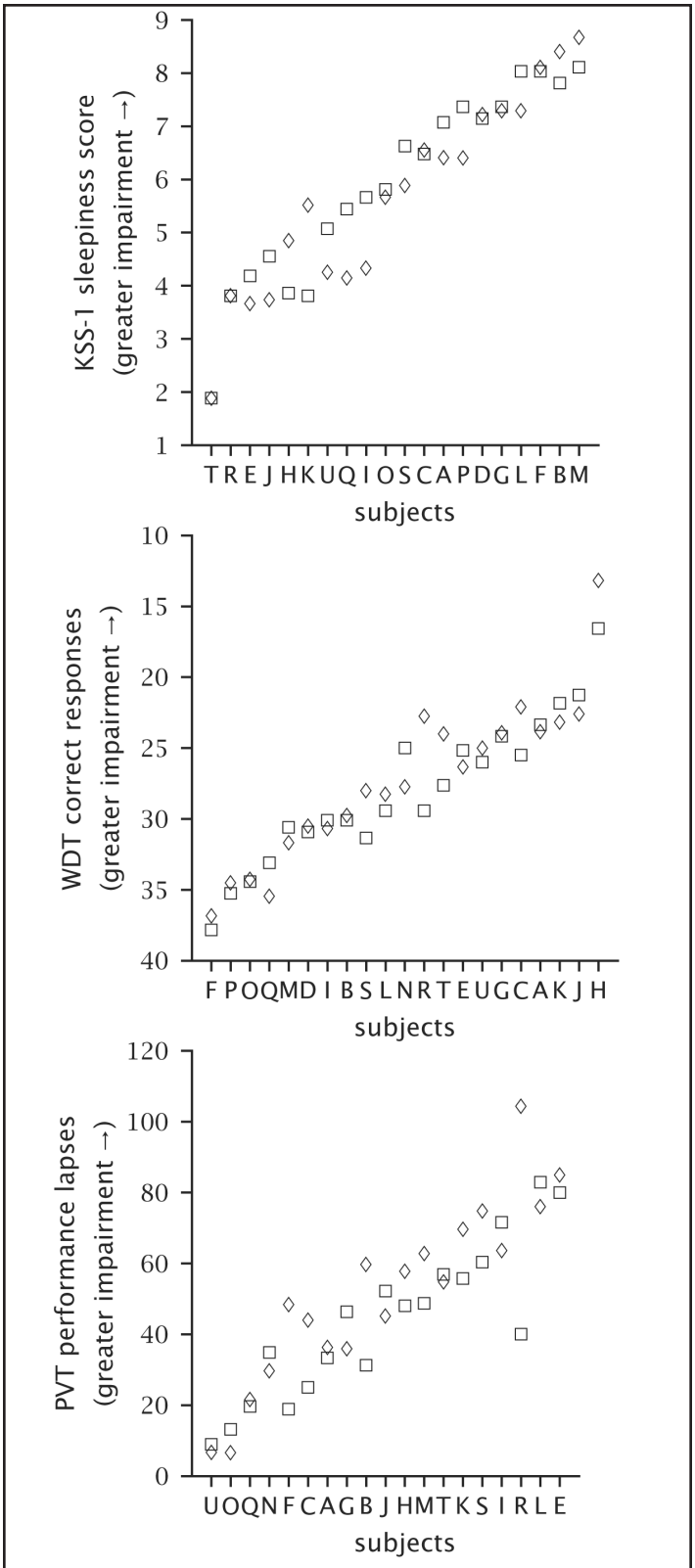

Figure 1—Neurobehavioral responses to total sleep deprivation, as determined by the averages over the last 24 hours of total sleep deprivation, in the 2 prior sleep extension (PSE) conditions. Data are shown for the first administration of the Karolinska Sleepiness Scale (KSS-1), the word detection task (WDT), and the psychomotor vigilance task (PVT). The abscissa of each panel shows the 21 individual subjects, labeled A through U, in arbitrary order with the same label being used for the same subject across the three panels. Within each panel, the subjects are ordered by the magnitude of their impairment (averaged over the 2 sleep deprivations), with the most resistant subjects on the left and the most vulnerable subjects on the right. Responses in the first exposure to sleep deprivation following 7 days of sleep extension are marked by boxes; responses in the second exposure to sleep deprivation following 7 days of sleep extension are marked by diamonds. KSS-1 data are missing for 1 subject, and PVT data are missing for 2 subjects. The panels reveal that subjects differed substantially in their responses to sleep deprivation, while the responses were relatively stable within subjects between the 2 exposures to sleep deprivation.

the within-subjects variance σ^2_{ws} in the data and computed the ICC. Table 2 shows the results of these analyses. Stable interindividual differences in the responses to TSD, as quantified by the ICC, were found to be statistically significant for all neurobehavioral tests ($Z > 2.3$, $P \leq .011$; see Table 2). Even with a Bonferroni correction for multiple comparisons, resulting in a type I error threshold of $\alpha = .004$, the ICC values would have been statistically significant for all neurobehavioral tests but one. According to the benchmarks defined by Landis and Koch,[38] the interindividual differences were "substantial" for the SAST, PVT, and PERF, and "almost perfect" (ie, highly robust) for all other neurobehavioral tests. Thus, under conditions of PSE (ie, after satiating any preexisting sleep debt), we found strong evidence for trait-like interindividual variability in neurobehavioral impairment from sleep loss. Order effects were statistically significant for the RARST, PVT, and VAS-M2 (see Table 2).

The analyses of trait-like interindividual variability were repeated with baseline daytime neurobehavioral functioning as a covariate. The results of these further analyses are shown in Table 3. Baseline differences contributed significantly to the between-subjects variability seen during TSD in all neurobehavioral tests ($t > 2.6$, $P < .02$; see Table 3). After controlling for daytime neurobehavioral functioning, no significant systematic interindividual differences remained during TSD for the SAST ($Z = 0.10$, $P = .46$) and the RARST ($Z < 0.01$, $P > .99$), and only a trend was found for the VAS-M1 ($Z = 1.64$, $P = .051$). On all other neurobehavioral tests, controlling for baseline functioning did not eliminate significant systematic interindividual variability (see Table 3), and ICC values continued to be indicative of trait-like interindividual differences in the responses to TSD. It is also noteworthy that no statistically significant order effects remained (see Table 3).

We observed that subjects' ranking in terms of the magnitude of impairment during TSD varied among neurobehavioral tests. That is, subjects showing the greatest deficits in one aspect of neurobehavioral functioning were not necessarily most impaired also in other aspects of neurobehavioral functioning (cf. Figure 1). Considering that the interindividual differences were robust (ie, high ICC values), we averaged each subject's neurobehavioral responses to TSD over the 2 PSE conditions (not correcting for baseline differences). A factor analysis was then performed to investigate whether or not interindividual differences were idiosyncratic to each neurobehavioral test. Inspection of the scree plot of eigenvalues revealed that 3 factors explained the larger part

Table 2—Results of variance components analyses performed to assess trait-like interindividual variability in impairment from sleep loss

Neurobehavioral Test	Variance Components					Order Effects	
	σ^2_{bs}	σ^2_{ws}	ICC	*Z*	*P*	*F*	*P*
KSS-1	3.92 (1.38)	0.42 (0.14)	0.904	2.84	.002	0.40	.81
VAS-F1	3.84 (1.36)	0.44 (0.15)	0.897	2.83	.002	0.39	.81
VAS-M1	1.59 (0.57)	0.21 (0.07)	0.882	2.79	.003	2.92	.053
SAST	495 (185)	138 (46)	0.782	2.68	.004	1.22	.34
DSST	2081 (760)	473 (158)	0.815	2.74	.003	2.22	.11
WDT	25.0 (8.5)	2.1 (0.7)	0.922	2.95	.002	2.99	.047
RARST	76.5 (27.5)	15.6 (5.2)	0.831	2.78	.003	2.25	.11
PVT	286 (124)	138 (48)	0.675	2.31	.011	4.36	.014
KSS-2	4.00 (1.44)	0.60 (0.21)	0.869	2.78	.003	0.51	.73
VAS-F2	4.21 (1.52)	0.67 (0.23)	0.862	2.76	.003	0.82	.53
VAS-M2	1.55 (0.55)	0.18 (0.06)	0.894	2.82	.002	3.28	.036
PERF	0.67 (0.25)	0.17 (0.06)	0.793	2.63	.004	1.78	.18
EFF	0.55 (0.19)	0.06 (0.02)	0.899	2.91	.002	1.37	.28

Using subjects' data from repeated exposure to total sleep deprivation under identical circumstances, the between-subjects variance σ^2_{bs} (a measure of systematic interindividual variability) was separated from the within-subjects variance σ^2_{ws} (a measure of variability between the 2 sleep deprivations); numbers in parentheses are estimated standard errors. For each of the neurobehavioral tests available for this study (see Table 1), the intraclass correlation coefficient (ICC) was computed from these variance components in order to quantify trait-like interindividual variability. Assessments were made of the statistical significance of this trait-like variability (*Z* statistics and *P* values are shown), as well as any order effects for which corrections were made (*F* statistics and *P* values are shown). Abbreviations are defined in Table 1.

(87%) of the variance in the data set. Table 4 shows the loadings of the different neurobehavioral tests on these 3 factors after orthogonal varimax rotation. A clustering of the neurobehavioral tests emerged on 3 distinct dimensions, which appeared to reflect self-evaluation of sleepiness, fatigue, and mood (factor 1); cognitive processing capability (factor 2); and behavioral alertness, as measured by the 20-minute PVT (factor 3). The self-evaluation dimension encompassed all 8 self-report measures of sleepiness, mood, effort, and performance. The cognitive processing dimension contained all objective measures of cognitive performance in the test battery except for the PVT, which was captured in the remaining dimension.

Correlates of Differential Vulnerability

Stepwise linear regression was used to investigate whether the interindividual differences during TSD were related to demographic parameters or sleep–wake-related variables. We found that pre-study overall sleepiness, as measured on the Epworth Sleepiness Scale,[27] was a significant predictor of the response to TSD as measured on the KSS-1 (explained variance: 27.1%), the KSS-2 (explained variance: 30.7%), and the VAS-F2 (explained variance: 27.0%). Pre-study sleepiness on the Epworth Sleepiness Scale and pre-study sleep quality on the Pittsburgh Sleep Quality Index[28] were both significant predictors of the response to TSD on the PERF rating (combined explained variance: 68.6%). The associations between these predictor variables and outcome variables were in the expected direction, with greater pre-study sleepiness and poorer pre-study sleep quality corresponding to greater deficits during TSD. The only other statistically significant finding in this analysis was that circadian phase (timing of the core body temperature minimum) predicted sleep-deprived performance on the SAST (explained variance: 28.5%), with later circadian phase corresponding to greater performance impairment.

Stepwise linear regression was also used to investigate whether the interindividual differences during sleep deprivation were related to psychosocial traits. Only the VAS-M2 had any statistically significant psychosocial predictors. The impatience subscale of the Survey of Work Styles[34] and the stressful conditions subscale of the Conflict-Stress Questionnaire[29] together explained 64.8% of the variance among subjects in this outcome variable; individuals who were more impatient but less easily stressed were relatively elated during TSD.

Effects of Sleep History

In order to assess the effect of the PSR condition relative to the 2 PSE conditions, we considered each subject's neurobehavioral responses to TSD for the different conditions simultaneously. Figure 2 illustrates these data, for the same neurobehavioral outcomes as shown in Figure 1. Using a mixed regression model, the magnitude β of the effect of sleep history was assessed simultaneously with the within-subjects variance σ^2_{ws} and the between-subjects variance σ^2_{bs} as well as any order effects. The results for σ^2_{bs}, β and order effects are shown in Table 5. For all neurobehavioral tests, trait-like interindividual differences in impairment from sleep deprivation were again statistically significant (see Wald Z tests of σ^2_{bs} in Table 5) and of similar magnitude as in the primary analyses of trait-like variability (cf. σ^2_{bs} in Table 2). Order effects were significant for the RARST and the PVT (see F tests in Table 5); RARST performance improved and PVT performance deteriorated with repeated exposure to TSD (independent of the order of the prior sleep conditions). The effect of sleep history β, corrected for order effects, was statistically significant for all neurobehavioral evaluations except the SAST, PVT, PERF, and EFF (see t tests in Table 5). All sleep-history effects reflected greater deficits during TSD with shorter sleep durations in the prior 7 days. Note that this involved positive β values for the SAST, DSST, WDT, and RARST (in which higher responses correspond to better performance), and negative β values for all other tests (in which lower responses correspond to better performance). The sleep-history analyses were repeated using only the total sleep times for the laboratory baseline nights, which yielded similar results.

For each neurobehavioral outcome variable, the population-average effect γ of the experimental manipulation of sleep history was determined by multiplying β (ie, the change in neurobehavioral functioning during TSD per additional hour of daily sleep obtained on average in the prior 7 days) by 4.1 hours (ie, the average difference in sleep duration between the PSR and the 2 PSE conditions). For easy comparison, Table 6 shows all statistically significant results for the population standard deviation σ_{bs} of neurobehavioral deficits during TSD, the population-average magnitude γ of the sleep history effect on neurobehavioral deficits during TSD, and the population-average changes from the first to the second and from the second to the third exposures to TSD. It is evident from the table that, within every neurobehavioral outcome, trait-like interindividual variability (σ_{bs}) was substantial in comparison with the magnitudes of the sleep-history effect (γ) and the order effects. On the basis of statistical distribution theory, we had derived that the effect of sleep history would still be a substantial determinant of neurobehavioral impairment from sleep loss—at least as important as the trait interindividual variability—if the criterion $|\gamma| / \sigma_{bs} \geq 0.68$ was met. The last column of Table 6 reveals that this was not the case for any of the neurobehavioral evaluations. Thus, the effect of sleep history as operationalized in this experiment was minor relative to trait interindividual variability, which was the more dominant determinant of subjects' vulnerability to neurobehavioral impairment from sleep loss.

Table 3—Results of variance components analyses for assessing trait-like interindividual variability in impairment from sleep loss, with baseline daytime neurobehavioral function as a covariate

Neurobehavioral Test	Variance Components					Order Effects		Baseline Covariate	
	σ^2_{bs}	σ^2_{ws}	ICC	*Z*	*P*	*F*	*P*	*t*	*P*
KSS-1	2.98 (1.06)	0.28 (0.10)	0.913	2.81	.003	0.84	0.52	3.83	.002
VAS-F1	2.35 (0.91)	0.39 (0.14)	0.857	2.59	.005	0.41	0.80	3.70	.002
VAS-M1	0.57 (0.35)	0.31 (0.14)	0.647	1.64	.051	1.22	0.34	4.39	.001
SAST	4 (39)	151 (51)	0.024	0.10	.46	0.26	0.90	10.70	<.001
DSST	476 (232)	369 (124)	0.563	2.05	.020	0.53	0.71	7.30	<.001
WDT	8.6 (4.0)	3.0 (1.2)	0.739	2.17	.015	1.13	0.38	5.01	<.001
RARST	<0.1*	22.3 (5.3)	<0.001	<0.01	>.99	0.78	0.56	10.92	<.001
PVT	190 (93)	121 (45)	0.611	2.05	.020	1.85	0.17	3.19	.006
KSS-2	2.21 (0.86)	0.41 (0.15)	0.844	2.56	.005	0.78	0.56	4.80	<.001
VAS-F2	1.98 (0.83)	0.61 (0.22)	0.763	2.37	.009	0.85	0.52	4.50	<.001
VAS-M2	0.52 (0.30)	0.23 (0.10)	0.694	1.70	.045	1.60	0.22	5.13	<.001
PERF	0.42 (0.17)	0.13 (0.05)	0.764	2.46	.007	2.63	0.073	4.11	.001
EFF	0.44 (0.16)	0.06 (0.02)	0.888	2.84	.002	1.22	0.34	2.69	.016

Details are the same as for Table 2. In addition, the statistical significance of the covariate was assessed (*t* statistics and *P* values are shown).
* Linearly dependent on the baseline covariate.

Table 4—Factor analysis results

Neurobehavioral Test	Factor 1	Factor 2	Factor 3
KSS-2	0.970	0.115	0.011
VAS-F2	0.963	0.193	–0.026
VAS-F1	0.958	0.147	0.002
KSS-1	0.938	–0.136	0.005
EFF	0.923	–0.130	–0.060
VAS-M1	0.892	0.222	–0.113
VAS-M2	0.808	0.316	–0.143
PERF	0.724	0.123	–0.315
RARST	0.253	0.908	–0.083
DSST	0.203	0.855	–0.223
WDT	–0.154	0.772	0.166
SAST	0.142	0.652	0.482
PVT	–0.116	0.004	0.672

Factor loadings of the 13 neurobehavioral tests (see Table 1) are shown, ordered by loading after orthogonal varimax rotation, for the 3 factors retained in the factor analysis.

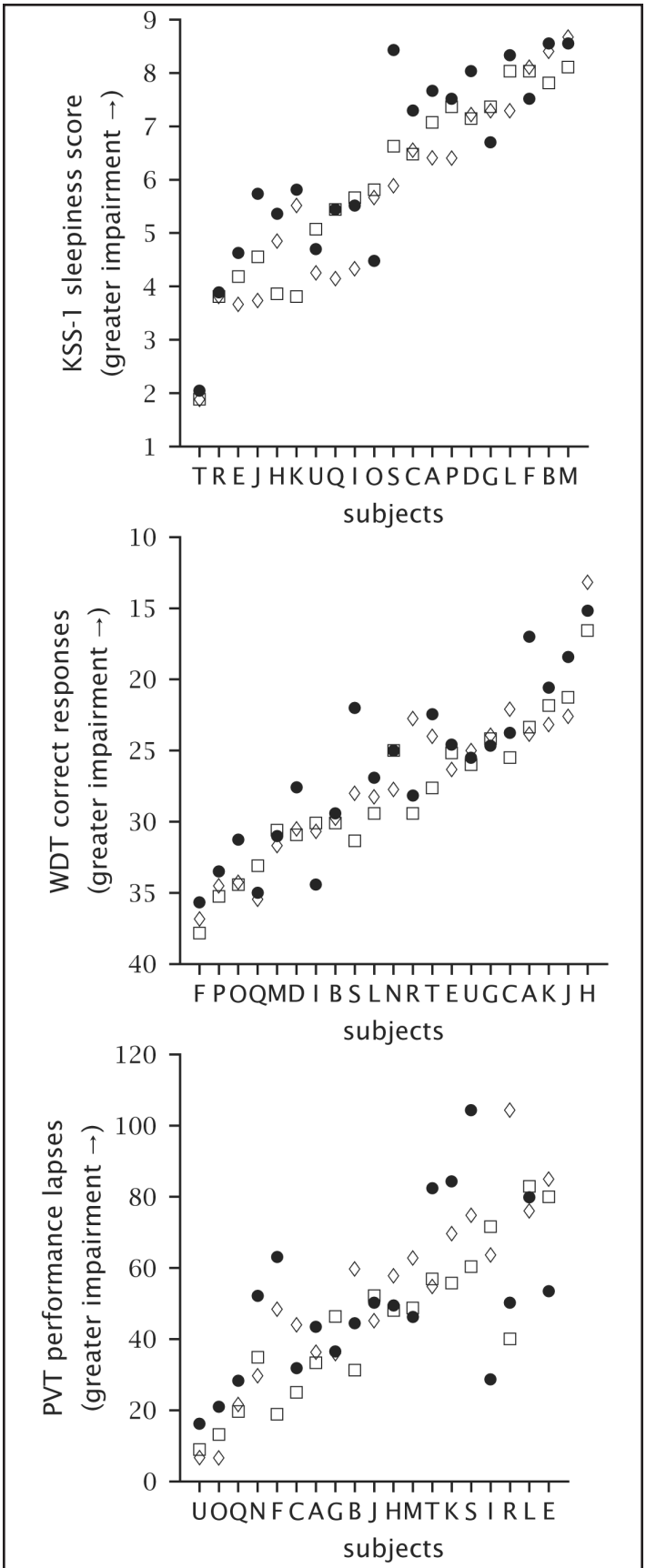

Figure 2—Neurobehavioral responses to sleep deprivation, as determined by the averages over the last 24 hours of sleep deprivation, in the 2 prior sleep extension (PSE) conditions and in the prior sleep restriction (PSR) condition. The panels correspond to those shown in Figure 1, with the responses to sleep deprivation after 7 days of sleep restriction (PSR condition) added as closed circles. The panels illustrate that subjects differed markedly in their overall responses to sleep deprivation, while the responses were relatively stable within subjects among the 3 exposures to sleep deprivation—regardless of sleep history. PVT refers to psychomotor vigilance task; WDT, word detection task; KSS, Karolinska Sleepiness Scale.

DISCUSSION

In this study involving repeated exposure to sleep deprivation under carefully controlled laboratory conditions, we found that neurobehavioral impairment from sleep loss was significantly different among individuals, stable within individuals, and robust relative to experimental manipulation of sleep history. Thus, this study is the first to demonstrate that interindividual differences in neurobehavioral deficits from sleep loss constitute a differential vulnerability trait.

In our sample of healthy young adults, substantial interindividual differences were observed during sleep deprivation for a variety of neurobehavioral tests commonly used in sleep-deprivation experiments. Across subjects, the responses to sleep loss covered wide intervals of the metric scales of these neurobehavioral outcomes (cf. Figure 1). Controlling for order effects, we quantified the interindividual differences by means of the ICC, which is a statistically valid measure of the proportion of variance in the data explained by systematic interindividual variability.[10] The ICC values showed that 67.5% to 92.2% of the variance in the neurobehavioral data was explained by stable variations among individuals (see Table 2). It should be noted that the ICC varies across populations depending upon their degree of homogeneity. In future studies, therefore, it is important to consider the magnitudes of between-subjects and within-subjects variance in addition to the ICC when comparing among populations. We reported the estimates of these variance components along with their standard errors (see Table 2) in order to facilitate comparisons across populations. Since our sample was comprised of a relatively homogeneous sample of healthy young adults, ICC values for the general population might be even higher than the ICC values reported here.

Although differences in baseline neurobehavioral functioning contributed significantly to the between-subjects variability during sleep deprivation, controlling for baseline functioning did not generally negate the systematic interindividual differences during sleep deprivation (see Table 3). Baseline performance did account for most of the interindividual differences observed during sleep deprivation in the SAST and the RARST, suggesting that variability in the response to sleep deprivation was dominated by baseline variability such as differences in aptitude for these specific tasks. Similarly, controlling for baseline differences cancelled out a portion of the interindividual variability during sleep deprivation in the VAS-M1, which may indicate that the mood response to sleep deprivation was affected by inherent differences in reporting bias or nonspecific differences in mood before sleep deprivation. These cases not withstanding, subjects showed differential vulnerability to neurobehavioral impairment during TSD across a range of objective and subjective tests.

It was hypothesized that interindividual variability in responses to sleep deprivation could be due to uncontrolled variations in sleep history. We manipulated sleep history experimentally by requiring subjects to either restrict their time in bed to 6 hours per day (PSR condition) or to extend their time in bed to 12 hours per day (PSE condition) in the 7 days prior to laboratory TSD. This procedure resulted in a difference of 4.1 hours in average daily sleep duration between the 2 conditions. The majority of neurobehavioral tests showed significant evidence of changes in performance during TSD due to this manipulation of sleep history (see Table 5). In all cases, however, these sleep-history effects were greatly exceeded by the magnitude of trait-like interindividual variability in the responses to sleep deprivation (see Table 6; cf. Figures 1 and 2). Thus, the hypothesis that interindividual variability in neurobehavioral deficits during sleep deprivation is primarily due to uncontrolled variations in sleep history must be rejected, with the caveat that more extreme variations in sleep history may yield greater contributions of this factor.

Order effects contributed significantly to the data from the RARST and the PVT (see Table 5). For the RARST, performance improved across the 3 sleep deprivations. This may reflect continuing learning across repetitive testing, which is typical for such cognitive performance tasks. For the 20-minute PVT, performance lapses increased over the

repeated exposures to TSD, even though there was ample opportunity (ie, at least 2 weeks) for recovery between sleep-deprivation sessions. We found no evidence that subjects became progressively more distressed by the PVT, and the VAS-M2 filled out after the PVT did not show any significant changes over the repeated sleep deprivations. Furthermore, there were no significant sleep-history and order effects on the EFF, indicating that subjects did not expend any differential amounts of effort among the 3 exposures to sleep deprivation. Order effects were quantitatively controlled for in all analyses of neurobehavioral responses, making it unlikely that order effects led to underestimation of the magnitude of the sleep-history effect. Nevertheless, the variance associated with order effects may have contributed to the statistical nonsignificance of the sleep-history effect for the PVT—in addition to a greater effect on psychomotor vigilance performance from acute TSD than from 7 days of (prior) sleep restriction.[4]

The trait-like interindividual differences in impairment from sleep loss were neither homogeneous across outcome measures nor idiosyncratic to the specific neurobehavioral tests. They clustered on 3 orthogonal dimensions that appeared to reflect self-evaluation of sleepiness, fatigue, and mood; cognitive processing capability; and behavioral alertness as measured by sustained attention (see Table 4). The self-evaluation assessments may have been orthogonal to the objective neurobehavioral tests because of their introspective nature or due to trait differences in subjective frame of reference or report bias. The emergence of more than 1 dimension for objective performance deficits indicates that distinct neurocognitive subsystems may regulate different aspects of the cognitive effects of sleep deprivation, as has also been suggested by recent neuroimaging findings.[39] This may have implications for understanding the neurobiology underlying the impairments resulting from sleep loss.

The existence of different dimensions to trait-like interindividual variability in neurobehavioral deficits from sleep loss may have consequences for operational settings that involve sleep deprivation,[40] in particular those where safety is an issue. The present results suggest that operational tasks depending on sustained attention (eg, monitoring of automated systems in factories, nuclear plants, defense systems, airplanes, etc.) may be affected by sleep loss in a different manner than are various other cognitive functions. Thus, individuals who are most at risk for loss of sustained attention due to sleep deprivation cannot be expected to surface when monitoring functioning on brief performance probes or fitness-for-duty tests involving cognitive processing. Moreover, individuals may subjectively judge themselves resistant to the effects of sleep loss, while neurobehavioral factors on which their work performance is based may actually be degraded. Therefore, as has been reported in other recent studies,[4,6,41] individuals cannot be relied upon to personally evaluate their performance capability and safety. Biomathematical models[42] may serve to predict neurobehavioral dysfunction more objectively, but these models should distinguish between different neurobehavioral functions as well. Available models must also first be expanded to allow for interindividual variability in predictions.[43] Recent progress in biomathematical modeling techniques[44] will facilitate such development.

Thus far, correlates of the trait responses to sleep deprivation have not been clearly identified. In a preliminary study,[45] we considered baseline sleep structure, but found only modest relationships between interindividual differences in baseline sleep architecture and neurobehavioral deficits during subsequent sleep deprivation. We also considered sleep need, even though interindividual differences in vulnerability to the effects of sleep loss are conceptually independent of interindividual differences in sleep need[3] (ie, long and short sleepers[46]). Using habitual sleep duration as a surrogate measure of sleep need in a stepwise linear regression, we observed no significant relationship with subjects' level of impairment during sleep deprivation. Thus, although the subject-selection criteria limited habitual sleep duration to between 6.5 and 8.5 hours, it would appear that differential vulnerability to neurobehavioral dysfunction during sleep deprivation is not merely a function of differential sleep need.

In addition to sleep variables, we considered circadian rhythm parameters as potential correlates of the trait responses to sleep deprivation, using core body temperature measurements during sleep deprivation to assess circadian rhythmicity. Circadian phase was predictive of performance deficits on the SAST but not on any other cognitive performance assay we used. This suggests that interindividual variability in circadian phase was not a reliable correlate of differential vulnerability to neuro-

Table 5—Results of mixed-model regression analyses for the effect of sleep history on neurobehavioral deficits during subsequent sleep deprivation

Neurobehavioral Test	Between-Subjects Variance			Order Effects		Effect of Sleep History		
	σ^2_{bs}	*Z*	*P*	*F*	*P*	β	*t*	*P*
KSS-1	3.78 (1.27)	2.96	.002	0.46	.63	–0.13 (0.04)	–3.03	.004
VAS-F1	3.83 (1.28)	2.99	.001	0.36	.70	–0.14 (0.04)	–3.75	.001
VAS-M1	1.74 (0.59)	2.92	.002	2.68	.082	–0.09 (0.03)	–2.65	.012
SAST	490 (168)	2.92	.002	0.74	.48	1.05 (0.67)	1.56	.13
DSST	1865 (637)	2.93	.002	3.12	.055	3.29 (1.30)	2.54	.015
WDT	27.3 (9.0)	3.03	.001	2.31	.11	0.33 (0.12)	2.88	.006
RARST	72.8 (24.2)	3.00	.001	6.37	.004	0.53 (0.21)	2.54	.015
PVT	355 (138)	2.57	.005	6.71	.003	–0.67 (0.86)	–0.78	.44
KSS-2	3.37 (1.15)	2.93	.002	1.20	.31	–0.15 (0.05)	–3.22	.003
VAS-F2	3.83 (1.30)	2.94	.002	1.03	.37	–0.12 (0.05)	–2.58	.014
VAS-M2	1.59 (0.56)	2.85	.002	2.25	.12	–0.09 (0.04)	–2.40	.021
PERF	0.64 (0.23)	2.81	.003	0.38	.69	–0.05 (0.03)	–2.00	.053
EFF	0.37 (0.13)	2.86	.002	1.99	.15	–0.01 (0.02)	–0.80	.43

For each of the available tests (see Table 1), the between-subjects variance σ^2_{bs} (with standard error) for neurobehavioral functioning during sleep deprivation (as a measure of systematic interindividual variability independent of sleep history) and the statistical significance thereof (*Z* statistics and *P* values) are shown. The table also displays the statistical significance (*F* statistics and *P* values) of order effects (independent of sleep history). Finally, the table shows the magnitude β (with standard error) of the effect of sleep history (expressed as the change in the neurobehavioral response to sleep deprivation for each additional hour of daily sleep obtained on average in the 7 days beforehand) and the statistical significance thereof (*t* statistics and *P* values).

Table 6—Magnitudes of interindividual variability, sleep-history effect, and order effects, for neurobehavioral impairment during sleep deprivation

Neurobehavioral Test	Interindividual Variability σ_{bs}	Sleep-History Effect γ	Order Effect (1st to 2nd)	Order Effect (2nd to 3rd)	$\lvert\gamma\rvert / \sigma_{bs}$
KSS-1	1.94	0.53	—	—	0.27
VAS-F1	1.96	0.58	—	—	0.30
VAS-M1	1.32	0.36	—	—	0.27
SAST	22.13	—	—	—	0.19
DSST	43.19	13.51	—	—	0.31
WDT	5.22	1.36	—	—	0.26
RARST	8.53	2.18	2.63	0.97	0.26
PVT	18.84	—	–9.46	–5.99	0.15
KSS-2	1.83	0.60	—	—	0.32
VAS-F2	1.96	0.49	—	—	0.25
VAS-M2	1.26	0.38	—	—	0.30
PERF	0.80	—	—	—	0.27
EFF	0.61	—	—	—	0.10

For direct comparison, the table shows the magnitude of trait-like interindividual differences σ_{bs} (square root of between-subjects variance σ^2_{bs} from Table 5); the magnitude of the effect of sleep history γ (the average difference in sleep duration between the prior sleep restriction and prior sleep extension conditions multiplied by the per-hour sleep-history effect β from Table 5); the magnitude of change from the first sleep deprivation session to the second (for neurobehavioral tests with significant order effects only); and the magnitude of change from the second sleep deprivation session to the third (*idem*). Every value in these columns is in the metric scale of the neurobehavioral test at hand (see Table 1). Metric scales for all neurobehavioral evaluations except the RARST, DSST, WDT, and SAST are reversed so that for each neurobehavioral evaluation, a positive sleep-history effect corresponds to greater deficits from total sleep deprivation (TSD) with prior sleep restriction compared to prior sleep extension, and a positive order effect corresponds to greater deficits in the earlier compared to the later exposure to TSD (— denotes not significant). The last column shows the absolute magnitude of the sleep-history effect expressed as a fraction of the magnitude of trait-like interindividual variability.

behavioral impairment from sleep loss in our study (note that subjects with extreme circadian phase preference were excluded from the sample). Masking effects from minor changes in posture and activity between performance test bouts may have affected our estimates of circadian parameters, however. Furthermore, any potential effects of interindividual differences in circadian rhythm parameters on sleep-deprived functioning in the two identical PSE conditions were limited by the design of the investigation, and it should be recalled that neurobehavioral performance outcomes were averaged over the circadian cycle.

We investigated a range of other potential correlates of the trait neurobehavioral responses to sleep deprivation, such as age, sex, and a variety of psychosocial factors. No systematic predictors of interindividual variability in the responses to sleep loss were found. Self-reported overall sleepiness as assessed before the study (Epworth Sleepiness Scale[27]) was an exception—it correlated with subjective evaluations of sleepiness and fatigue during sleep deprivation (although little variance was explained). This result would seem to provide some support for the hypothesized concept of trait (subjective) sleepiness.[47] Even so, it appears that correlates and biomarkers of interindividual variability in objectively measured cognitive vulnerability to sleep deprivation have yet to be discovered.

Although the present documentation of trait-like interindividual variability may complicate the analysis of data from sleep deprivation studies and other experimental or clinical interventions, taking this variability into consideration is important. The phenotypic variability in neurobehavioral responses to sleep loss is substantial; in a recent study of chronic sleep restriction,[4] for example, systematic interindividual differences accounted for more than 60% of the variance in the data. Investigations based solely on population-average responses—as is typical throughout the field of sleep research—may produce misleading conclusions.[10,44,48] In another study of chronic sleep restriction,[5] it was reported that recuperation from the neurobehavioral deficits accumulated during 7 days of partial sleep deprivation was still incomplete after 3 days with recovery sleep. Since interindividual differences were not taken into account in the analyses for that study, it remains unclear whether the apparent lack of recovery was a general phenomenon or was restricted to a few subjects whose responses could not be balanced by the rest of the subjects (due to the floor effect associated with full recovery). Such limitations can be overcome with mixed-model statistical approaches[36,37,48] like those used in the present study and in our recent investigation of chronic sleep restriction.[4] These techniques allow for simultaneous estimation of the population-mean effect of an experimental intervention, as well as the variability among subjects in the response to the intervention. This may be helpful, for instance, to understand the reported poor relationship between severity of obstructive sleep apnea and severity of daytime impairment,[49] which could be due to interindividual differences in vulnerability to the effects of sleep loss.

Given the large portion of the variance in cognitive and psychomotor vigilance performance explained by stable interindividual differences in responses to sleep deprivation, there is a critical need to identify the neurobiologic basis of trait-like differential vulnerability. Functional genomics and proteomics, functional neuroimaging approaches, and neuropharmacologic studies may reveal mechanisms subserving these phenotypic responses. Since the function or functions of sleep are still actively debated,[50,51] focusing on interindividual differences in responses to sleep deprivation and the relationship, if any, with sleep architecture may also provide new insights into the question of why we need sleep.[52,53]

Finally, we note that while the TSD sessions of our study were performed inside a laboratory in a clinical research center, the manipulation of sleep history relied on subjects restricting their sleep at home for 6 of the 7 sleep-restriction condition days. In the PSR condition, therefore, subjects were instructed explicitly to not engage in safety-sensitive tasks such as driving or operating heavy machinery. Nevertheless, in light of our recent findings on the cumulative neurocognitive effects of chronic restriction of sleep to 6 hours per day,[4] we no longer consider it safe to engage in such moderate sleep restriction outside the controlled environment of the laboratory. Thus, a follow-up study currently underway is being conducted entirely inside the laboratory so as to ensure subjects' safety when they are deprived of sleep.

ACKNOWLEDGMENTS

We thank the subject volunteers; the laboratory staff, technicians, and student behavioral monitors; physicians Martin Szuba, MD and Christos Ballas, MD; the staff of the General Clinical Research Center of the Hospital of the University of Pennsylvania; and Els Møst, Allison Stakofsky, Marjolijn Blaauw, Liang-tso Tung, Jennifer Parker, Daniel Green, Alicia Levin, Geoffrey Nosker, Claire Fox, and John Powell for help with data reduction and analysis. This work was supported by NASA grant NAG9-1161 and NIH grants R01-HL70154 and HL70154-02S1 awarded to Hans Van Dongen, and NIH grant M01-RR00040 awarded to the General Clinical Research Center of the Hospital of the University of Pennsylvania.

REFERENCES

1. Dinges DF, Kribbs NB. Performing while sleepy: effects of experimentally-induced sleepiness. In: Monk TH, ed. Sleep, sleepiness and performance. Chichester: John Wiley & Sons, 1991:97-128.
2. Harrison Y, Horne JA. Sleep loss impairs short and novel language tasks having a prefrontal focus. J Sleep Res 1998;7:95-100.
3. Van Dongen HPA, Rogers NL, Dinges DF. Sleep debt: theoretical and empirical issues. Sleep Biol Rhythms 2003;1:5-13.
4. Van Dongen HPA, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 2003;26:117-28.
5. Belenky G, Wesensten NJ, Thorne DR, et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. J Sleep Res 2003;12:1-12.
6. Leproult R, Colecchia EF, Berardi AM, Stickgold R, Kosslyn SM, Van Cauter E. Individual differences in subjective and objective alertness during sleep deprivation are stable and unrelated. Am J Physiol Regulat Integr Comp Physiol 2003;284:R280-90.
7. Wilkinson RT. Interaction of lack of sleep with knowledge of results, repeated testing, and individual differences. J Exp Psychol 1961;62:263-71.
8. Webb WB, Levy CM. Effects of spaced and repeated total sleep deprivation. Ergonomics 1984;27:45-58.
9. Van Dongen HPA, Dijkman MV, Maislin G, Dinges DF. Phenotypic aspect of vigilance decrement during sleep deprivation. Physiologist 1999;42:A-5.
10. Van Dongen HPA, Maislin G, Dinges DF. Dealing with interindividual differences in the temporal dynamics of fatigue and performance: importance and techniques. Aviat Space Environ Med 2004;75:A147-54.
11. Cox DR, Solomon PJ. Components of Variance. Boca Raton: Chapman & Hall/CRC; 2003.
12. Dijkman M, Sachs N, Levine E, et al. Effects of reduced stimulation on neurobehavioral alertness depend on circadian phase during human sleep deprivation. Sleep Res 1997;26:265.
13. Smith CS, Reilly D, Midkiff K. Evaluation of three circadian rhythm questionnaires with suggestions for an improved measure of morningness. J Appl Psychol 1989;74:728-38.
14. Douglass AB, Bornstein R, Nino-Murcia G, et al. The Sleep Disorders Questionnaire I: Creation and multivariate structure of SDQ. Sleep 1994;17:160-7.
15. Åkerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. Int J Neurosci 1990;52:29-37.
16. Monk TH, Embrey DE. A visual analogue scale technique to measure global vigor and affect. Psychiatry Res 1989;27:89-99.
17. Thorne DR, Genser SG, Sing GC, Hegge FW. The Walter Reed performance assessment battery. Neurobehav Toxicol Teratol 1985;7:415-8.
18. Wechsler D. Wechsler Adult Intelligence Scale-Revised. New York: Psychological Corp; 1981.
19. Jex HR, McDonald JD, Phatak AV. A "critical" tracking task for man-machine research related to the operator's effective delay time. I. Theory and experiments with a first-order divergent controlled element. NASA contractor report 616; 1966.
20. Bickel WK, Higgins ST, Griffiths RR. Repeated diazepam administration: Effects on the acquisition and performance of response chains in humans. J Exp Anal Behav 1989;52:47-56.
21. Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behav Res Meth Instr Comp 1985;17:652-5.
22. Dinges DF, Kribbs NB, Steinberg KN, Powell JW. Do we lose the willingness to perform during sleep deprivation? Sleep Res 1992;21:318.
23. Rechtschaffen A, Kales A, eds. A Manual of Standardized Terminology, Techniques And Scoring System for Sleep Stages of Human Subjects. Los Angeles: UCLA Brain Information Service; 1968.
24. Brown EN, Czeisler CA. The statistical analysis of circadian phase and amplitude in constant-routine core-temperature data. J Biol Rhythms 1992;7:177-202.
25. Kerkhof GA, Van Dongen HPA. Morning-type and evening-type individuals differ in the phase position of their endogenous circadian oscillator. Neurosci Lett 1996;218:153-6.
26. Van Dongen HPA, Kerkhof GA, Souverijn JHM. Absence of seasonal variation in the phase of the endogenous circadian rhythm in humans. Chronobiol Int 1998;15:623-32.

27. Johns MW. Sleepiness in different situations measured by the Epworth Sleepiness Scale. Sleep 1994;17:703-10.
28. Buysse DJ, Reynolds III CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193-213.
29. Steinmetz JI, Kaplan RM, Miller GL. Stress management: An assessment questionnaire for evaluating interventions and comparing groups. J Occup Med 1982;24:923-31.
30. Eysenck HJ. Biological Basis of Personality. Springfield: Thomas; 1967.
31. Crowne DP, Marlowe D. The Approval Motive. New York: Wiley, 1964.
32. Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the Penn State Worry Questionnaire. Behav Res Ther 1990;28:487-95.
33. Barker BM, Barker Jr HR, Wadsworth Jr AP. Factor analysis of the items of the state-trait anxiety inventory. J Clin Psychol 1977;33:450-5.
34. Jackson DN, Mavrogiannis A. An Introduction to the Survey of Work Styles. University of Western Ontario research bulletin 663; 1987.
35. Myers IB. Introduction to Type. Palo Alto: Consulting Psychologists Press; 1998.
36. Burton P, Gurrin L, Sly P. Extending the simple linear regression model to account for correlated responses: An introduction to generalized estimating equations and multi-level mixed modelling. Stat Med 1998;17:1261-91.
37. Van Dongen HPA, Olofsen E, Dinges DF, Maislin G. Mixed-model regression analysis and dealing with interindividual differences. In: Johnson ML, Brand L, eds. Numerical Computer Methods (part E). San Diego: Academic Press 2004:139-171.
38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
39. Drummond SPA, Brown GG. The effects of total sleep deprivation on cerebral responses to cognitive performance. Neuropsychopharmacol 2001;25:S68-73.
40. Gillberg M, Åkerstedt T. Individual differences in susceptibility to sleep loss. In Haider M, Koller M, Cervinka R, eds. Proceedings of the VII International Symposium on Night- and Shiftwork, 1985:117-22.
41. Dorrian J, Lamond N, Holmes AL, et al. The ability to self-monitor performance during a week of simulated night shifts. Sleep 2003;26:871-7.
42. Mallis MM, Mejdal S, Nguyen TT, Dinges DF. Summary of the key features of seven models of human fatigue and performance. Aviat Space Environ Med 2004;75:A4-14.
43. Van Dongen HPA, Balkin TJ. Commentary on modeling performance and alertness: the QinetiQ approach. Aviat Space Environ Med 2004;75:A104-5.
44. Olofsen E, Dinges DF, Van Dongen HPA. Nonlinear mixed-effects modeling: Individualization and prediction. Aviat Space Environ Med 2004;75:A134-40.
45. Møst EIS, Baynard MD, Fox CG, et al. On the relationship between interindividual differences in performance impairment from sleep loss and interindividual differences in sleep architecture. Sleep-Wake Research in the Netherlands 2004;14:55-8.
46. Aeschbach D, Cajochen C, Landolt H, Borbély AA. Homeostatic sleep regulation in habitual short sleepers and long sleepers. Am J Physiol 1996;270:R41-53.
47. Cluydts R, De Valck E, Verstraeten E, Theys P. Daytime sleepiness and its evaluation. Sleep Med Rev 2002;6:83-96.
48. Feldman HA. Families of lines: random effects in linear regression analysis. J Appl Physiol 1988;64:1721-32.
49. Kribbs NB, Getsy JE, Dinges DF. Investigation and management of daytime sleepiness in sleep apnea. In: Saunders NA, Sullivan CE, eds. Sleep and Breathing. New York: Marcel Dekker, 1994:575-604.
50. Stickgold R, Hobson JA, Fosse R, Fosse M. Sleep, learning, and dreams: off-line memory reprocessing. Science 2001;294:1052-7.
51. Siegel JM. The REM sleep-memory consolidation hypothesis. Science 2001;294:1058-63.
52. Horne J. Why sleep? Biologist 2002;49:213-6.
53. Siegel JM. Why we sleep. Sci Am 2003;289:92-7.