

Random number generation during sleep deprivation: effects of caffeine on response maintenance and stereotypy

JULIE M. GOTTSSELIG, MARTIN ADAM, JULIA V. RÉTEY, RAMIN KHATAMI, PETER ACHERMANN and HANS-PETER LANDOLT

Section of Psychopharmacology and Sleep Research, Institute of Pharmacology and Toxicology, University of Zurich, Zurich, Switzerland

Accepted in revised form 2 October 2005; received 13 April 2005

SUMMARY Neurophysiological and functional imaging studies have demonstrated that frontal regions of the brain are particularly responsive to homeostatic sleep pressure. Previous neuropsychological studies indicate that sleep deprivation causes impairments in prefrontal cortical function. Random number generation (RNG) is thought to provide a sensitive index of executive functions that rely on the prefrontal cortex. The present study tested the hypothesis that sleep deprivation would impair RNG and that caffeine would mitigate this impairment. Healthy young men ($n = 21$) participated in two 40-h sleep deprivations 1 week apart. During each sleep deprivation period subjects received either caffeine or placebo according to a randomized, double-blind cross-over design, and they completed an oral RNG task at 3-h intervals. Comparison of test sessions at analogous times of day revealed that sleep deprivation was associated with significant drops in the number of responses, a threefold increase in the percentage of rule violations, 59% greater response redundancy and a 20% increase in stereotypy of adjacent response pairs. Sleep deprivation did not consistently alter counting tendency. Caffeine ameliorated the decrease in the number of responses but did not mitigate other deficits in RNG that arose during sleep deprivation. These findings are consistent with prior reports of diminished vigilance and increased perseveration during extended wakefulness. They support the conclusion that caffeine preserves simple aspects of cognitive performance during sleep deprivation, whereas caffeine may not prevent detrimental effects of sleep deprivation on some complex cognitive functions.

KEYWORDS executive function, frontal cortex, neurobehavioral performance, perseveration, random numbers, sleepiness, stimulants

INTRODUCTION

Random number generation (RNG) tests provide a sensitive measure of frontal executive function (Baddeley *et al.*, 1998; Brugger *et al.*, 1996). Frontal lesions impair random generation (Spatt and Goldenberg, 1993; Wiegiersma *et al.*, 1990), and improvements in random generation correlated with clinical improvements among psychiatric inpatients (Horne

et al., 1982). Neuroimaging studies indicate involvement of the frontal cortex in the performance of RNG tasks (Artiges *et al.*, 2000; Itagaki *et al.*, 1995; Jahanshahi *et al.*, 2000). Activation of the superior anterior cingulate correlated with greater randomness in the choice of neighboring response pairs (Artiges *et al.*, 2000), and activation of the left dorsolateral prefrontal cortex was associated with suppression of habitual counting responses (Jahanshahi *et al.*, 2000). When subjects randomly generated response choices in either of two response modalities, increased blood flow was observed in the dorsolateral prefrontal cortex (Frith *et al.*, 1991). Studies with transcranial magnetic stimulation indicated that in random number and letter generation tasks, the

Correspondence: J.M. Gottselig, Division of Sleep Medicine, Harvard Medical School, Brigham and Women's Hospital, 221 Longwood Ave., Boston, MA 02115, USA. Tel.: (617) 732 5206; fax: (617) 732 4015; e-mail: jgottselig@rics.bwh.harvard.edu

left dorsolateral prefrontal cortex serves to inhibit habitual responses (Jahanshahi and Dirnberger, 1999; Jahanshahi *et al.*, 1998).

Frontal regions of the brain appear to be particularly sensitive to homeostatic sleep pressure (Horne, 1993). Sleep electroencephalogram (EEG) slow-wave activity (approximately 0.5–4.5 Hz) indexes sleep pressure (Borbély and Achermann, 2005) and is prominent in frontal EEG derivations (Buchsbaum *et al.*, 1982; Werth *et al.*, 1997). After sleep deprivation, sleep EEG slow-wave activity increases, and the increase in slow-wave activity is most pronounced in frontal EEG derivations (Cajochen *et al.*, 1999; Finelli *et al.*, 2001). Fluorodeoxyglucose positron emission tomography imaging (FDG-PET) studies of subjects who performed a serial addition/subtraction task in baseline conditions and after 24 h of wakefulness revealed that sleep deprivation reduced regional cerebral metabolic rate in the prefrontal cortex (Thomas *et al.*, 2000). Functional magnetic resonance imaging studies similarly revealed that during performance of serial subtraction prefrontal cortical activation decreased after 35 h of wakefulness (Drummond *et al.*, 1999); in contrast, during performance of verbal learning and divided attention prefrontal activation increased (Drummond *et al.*, 2001). The increases were associated with increases in subjective sleepiness and may represent cerebral compensation for the effects of sleep deprivation (Drummond *et al.*, 2000). People who are sleep deprived exhibit cognitive deficits that resemble those observed in patients with prefrontal cortical lesions. These deficits include increased perseveration (Harrison and Horne, 1997, 1999; Horne, 1988), impaired planning abilities (Harrison and Horne, 1999, 2000a) and flatness of speech (Harrison and Horne, 1997). Recent evidence suggests that local prefrontal sleep EEG rhythms are associated with daytime performance (Anderson and Horne, 2003).

Based on prior evidence that RNG is sensitive to deficits in frontal executive function and that sleep deprivation alters prefrontal cortical function, we predicted that sleep deprivation would impair RNG performance. Sleep deprivation and prefrontal lesions have been associated with stereotyped responding (perseveration). Thus, we hypothesized that sleep deprivation would increase two measures of response stereotypy: (1) redundancy, which quantifies the extent to which some response alternatives are employed more frequently than others and (2) the null score quotient, which quantifies stereotypy in adjacent response pairs. Furthermore, because sleep deprivation may impair the functioning of prefrontal areas that are involved in the suppression of counting tendency, we hypothesized that sleep deprivation would increase counting tendency (cf. Heuer *et al.*, 2005).

We investigated the effects of sleep deprivation on RNG within the context of a double-blind cross-over study of the effects of caffeine during sleep deprivation (Landolt *et al.*, 2004). Previous studies of the effects of caffeine on performance during sleep deprivation have focused primarily on measures of simple cognitive processes, and the potential

benefit of caffeine on measures of executive function merits investigation (Wyatt *et al.*, 2004). It was expected that if sleep deprivation impaired RNG, caffeine would reduce this impairment.

METHODS

Subjects

Participants were male students [$n = 23$, aged 20–30 years, mean = 24.7 ± 0.6 (SEM) years] recruited from the University of Zürich and the Swiss Federal Institute of Technology and paid for their participation in the study. All subjects were non-smokers who reported that they were in good health, had no history of neurologic or psychiatric disease and had not taken any medications or consumed any illicit drugs in the 2 months before the study. They reported habitual alcohol consumption of less than seven drinks per week and habitual caffeine consumption of less than 300 mg per day. They indicated that they were good sleepers with regular bedtimes and no subjective sleep disturbances. Based on polysomnographic screening, subjects with sleep apnea, sleep efficiency < 75% or a periodic leg movements in sleep (PLMS) index of 5 or more per hour of sleep were excluded from participation. The participants in the present study include those in the report of Landolt *et al.* (2004). The data of two subjects were excluded from analyses because these subjects did not comply with the task instructions. Study procedures were approved by the local ethics committee, and subjects gave written informed consent before participating in the study.

Pre-experimental procedure

For 2 weeks prior to the study, subjects were required to abstain from all dietary sources of caffeine, wear a wrist activity monitor on the non-dominant arm and keep a sleep-wake diary. During the 3 days prior to the study, participants were required to abstain from alcohol intake and to keep a regular 8-h night-time sleep schedule from either 23:00 to 07:00 hours ($n = 4$) or from 00:00 to 08:00 hours ($n = 17$). These times corresponded approximately to the subjects' habitual sleep times, and subjects were not allowed to deviate from these times by more than 1 h. Compliance with these requirements was verified with daily log books and wrist actigraphy. Upon entering the laboratory, salivary caffeine assays and breathalyzer tests were conducted to confirm participants' abstinence.

Procedure

The study design is illustrated in Fig. 1. Subjects completed a RNG task 14 times at 3-h intervals during the course of two 40-h sleep deprivation periods that occurred 1 week apart. Subjects received 200 mg caffeine, 11 and 23 h into one of the sleep deprivation periods. [One cup or 225 mL of ground roast 'brewed' coffee has about 125 mg of caffeine; however, the

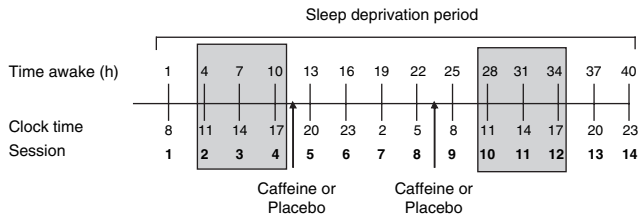


Figure 1. Experimental protocol. Subjects underwent two 40-h sleep deprivation periods that occurred 1 week apart. During one of the periods subjects received caffeine at 11 and 23 h into the sleep deprivation (200 mg at each time) and during the other they received placebo at these times, according to a randomized, double-blind cross-over design. The random number generation task was administered every 3 h beginning 30–40 min after awakening (the ‘time awake’ given in this and subsequent figures is rounded up to the nearest hour). The clock times on this and subsequent figures are given in military time (e.g. such that 14 represents 2 PM, etc.) and represent the midpoint of the times when the task was administered to subjects who slept from 11 PM to 7 AM ($n = 4$, task administration beginning at 7:30 AM) and those who slept from 12 AM to 8 AM ($n = 17$, task administration beginning at 8:30 AM). The gray shading indicates sessions at analogous times of day; data from these sessions were compared to assess the effects of sleep deprivation.

caffeine content of a cup of coffee varies widely depending on the preparation method, type of coffee bean, etc. (James, 1997).] During the other sleep deprivation period subjects received placebo, according to a randomized, double-blind cross-over design. Prior to each sleep deprivation period subjects spent two nights in the laboratory. On these nights the subjects had an 8-h sleep opportunity scheduled at the accustomed time. On the evening of the first night in the laboratory, they completed the RNG task once to familiarize themselves with the task. A previous study suggested that performance in RNG remains stable with repeated test administration (Evans and Graham, 1980).

Random number generation task

Subjects orally generated sequences of numbers using the response alternatives 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. They were instructed to generate numbers in a sequence that should be as random as possible, such as a sequence which would result if one were to throw a 10-sided die to generate each number. The task required generation of 225 numbers at each of two paces: slow (one number every 1800 ms) and fast (one number every 750 ms). The pace was set by a tone generated by a computer. The slow portion of the task always preceded the fast portion, and there was a short break between these portions. The task lasted approximately 10 min. Responses were recorded on dictation machines and transcribed at a later date.

Data analysis

Dependent variables

Sample size deviation indicates deviations from the requested sample size. It is equal to $n-225$. The sample size, n , is the

number of numbers that a subject generated. It does not include responses that were not allowed. If subjects made a response that was not allowed (e.g. if they used a number greater than 9 or made a response that was not a number), the response was transcribed as a rule break.

Rule breaks were quantified as the percentage of responses that were not allowed. For each RNG test, the number of rule breaks was divided by the sample size and multiplied by 100.

Redundancy is a measure of zero-order response stereotypy that quantifies the extent to which subjects used some response alternatives more frequently than others. It is a measure of deviation from maximum information generation and was calculated according to the following formula (Towse and Neil, 1998):

$$\text{Redundancy} = 100 \times \left(1 - \frac{\log_2 n - n^{-1} \sum_i^a n_i \log_2 n_i}{\log_2 a} \right)$$

where n is the sample size, n_i the number of occurrences of the i th response alternative and a the number of response alternatives.

The *null score quotient* (NSQ) is a measure of first-order response stereotypy that quantifies stereotypy in adjacent response pairs. It indicates the percentage of possible response pairs (digrams) that were not used in the response set (Towse and Neil, 1998):

$$\text{NSQ} = 100 \times \left(\frac{d}{a^2 - 1} \right)$$

where a is the number of response alternatives and d the number of digram permutations that do not appear in the response set. The value a^2 gives the number of possible digram permutations. The value of d can vary from ‘0’ to a^2-1 . To get the value that is used in the denominator, a ‘1’ must be subtracted from the a^2 because if subjects supplied a response sequence, they must have used at least one of the possible digrams.

Adjacency quantifies counting tendency. It was calculated as the percentage of response pairs within a sequence for which the numbers were adjacent items from an ordinal sequence of alternatives (Towse and Neil, 1998). Response pairs of ‘2, 3’ or ‘9, 8’ are examples of adjacent responses.

$$\text{Adjacency} = 100 \times \frac{\text{number of adjacent pairs}}{\text{number of response pairs}}$$

Normalization

During the RNG task, subjects sometimes did not speak precisely at the rate of the pacing tone, which resulted in variations in the length of the number sequences. To permit comparison of results derived from number sequences of different lengths, the redundancy, NSQ and adjacency scores were normalized with respect to computer-generated random sequences. For each sequence length empirically observed in our dataset, we computed 1000 computer-generated random sequences of the corresponding length using a Matlab implementation (by Christian Merkwirth) of the Mersenne Twister method (Matsumoto and Nishimura, 1998). The dependent

variables were derived from each of these computer-generated random sequences. Then, for each sequence length, the means and standard deviations of each dependent variable were calculated over these 1000 values. Z-scores were calculated for the dependent variables derived from the experimental data by subtracting the means and dividing by the standard deviations of the corresponding dependent variables derived from computer-generated data that had the same sequence length as the human-generated data. Prior to the normalization all redundancy data were log-transformed so that the data were normally distributed. Calculation and normalization of dependent variables were conducted using Matlab version 6.5 (The MathWorks, Natick, MA, USA).

Statistics

With the exception of the rule breaks, the dependent variables were analyzed using covariance pattern models that were four-way analyses of variance with the factors condition (caffeine and placebo), speed (fast and slow), week (1 and 2), session (1–14) and all interactions of these factors. Statistically significant results ($P < 0.05$) are reported. As is conventional, the error bars shown in the figures represent the between-subjects variability.

The mixed model ANOVAs were calculated using SAS proc mixed (SAS Version 9.1, SAS Institute Inc., Cary, NC, USA). Mixed models have the advantage that they permit analysis of repeated measures data when data are missing and they allow specification of different types of covariance structures. Comparison of different covariance structures makes it possible to take into account the correlations between observations collected from the same subject. The present study involved repeated measures observations (in each of 14 sessions) taken in each of the two laboratory visits (week 1 or week 2). Several covariance structures were tried for each dependent variable (Brown and Prescott, 1999), and the model with the best fit was chosen based on information criteria (Akaike's Information Criterion and Schwarz's Bayesian Criterion) and on the residual log likelihood values, which permitted statistical comparisons between models that subsumed each other. If the models fitted did not differ significantly according to a chi-square comparison, the simpler model (i.e. that involving fewer covariance parameters) was employed. The restricted maximum likelihood (REML) method for variance component estimation was employed, and the Satterthwaite method was used for computing the denominator degrees of freedom for the tests of fixed effects. If changes across sessions were significant, the SAS 'estimate' statement was used to compare results from sessions that occurred at analogous times of day (11:00, 14:00 and 17:00) before and after the night of sleep deprivation. Thus, the mean of sessions 2–4 was compared with the mean of sessions 10–12 (see Fig. 1). Sessions 1 and 9 were not included in this comparison because session 1 occurred soon after awakening and may have been influenced by sleep inertia.

Because the median percentages of rule breaks did not meet the assumptions required for parametric testing, they were

analyzed using non-parametric Wilcoxon matched-pairs signed-ranks tests calculated using the software JMP (SAS Institute Inc., Cary, NC, USA). Tests of the effects of condition, week, speed, and session were conducted. Two-way interactions were tested by performing multiple tests of one independent variable at all levels of the other independent variable. To correct for multiple comparisons, a Bonferroni-corrected alpha criterion was employed. For example, to test the speed \times session interaction, tests of the effect of speed were conducted for each session, and an α -criterion $0.05/14 = 0.0037$ was used.

RESULTS

Sample size deviation

As shown in Fig. 2, subjects generally tended to generate fewer numbers than requested, and the negative deviation became more pronounced with increasing sleep deprivation [session effect: $F(13, 248) = 3.32$, $P = 0.0001$]. There was a significant session \times condition interaction [$F(13, 776) = 4.48$, $P < 0.0001$], because caffeine reduced the negative deviation on the second day of the sleep deprivation. Comparison of baseline and sleep deprivation sessions at analogous times of day (11:00, 14:00 and 17:00) revealed that after sleep deprivation there was a significant decrement in the sample size in the placebo condition (estimated difference = -8.8 ± 2.1 , $P < 0.0001$), whereas there was no significant decrement in the caffeine condition (estimated difference = -0.41 ± 1.6 ,

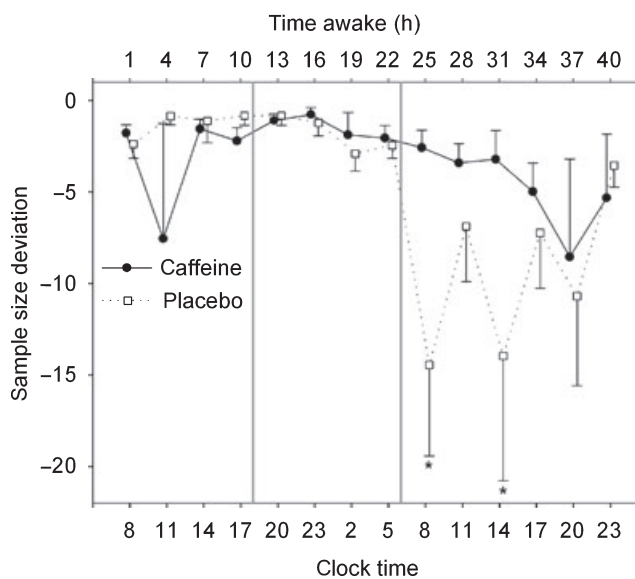


Figure 2. Sample size deviation, which represents the difference in the number of numbers subjects generated from the number of numbers required by the task (i.e. the observed sample size–225). The asterisks indicate significant differences between the conditions ($P < 0.001$, Tukey–Kramer *post hoc* tests). The vertical lines indicate the times of caffeine or placebo administration. The error bars represent standard errors. Values are staggered along the x-axis so that it is easier to see the error bars.

$P = 0.80$). The estimated negative deviation associated with sleep deprivation was 21 times greater in the placebo condition than in the caffeine condition, and the difference between the caffeine and placebo conditions was statistically significant ($P = 0.04$).

Rule breaks

The median percentage of rule breaks changed significantly across sessions (see Fig. 3; $\chi^2 = 47.8$, $P < 0.0001$). Comparison of sessions at analogous times of day revealed that rule breaks increased by more than 3-fold after the night of sleep deprivation (difference in the median percentage of rule breaks = $0.49 \pm 0.14\%$, $P < 0.001$ based on Wilcoxon signed-rank test; means and standard errors before: $0.22 \pm 0.10\%$, after: $0.71 \pm 0.16\%$). The median percentage of rule breaks was greater for the fast than for the slow condition in week 1 (difference = $0.24 \pm 0.06\%$, $P = 0.002$; fast: $0.47 \pm 0.10\%$, slow: $0.24 \pm 0.08\%$), but not in week 2 (difference = $0.17 \pm 0.13\%$, $P = 0.30$; fast: $0.37 \pm 0.14\%$, slow: $0.20 \pm 0.07\%$). There was a greater percentage of rule breaks for the fast than the slow speed in session 12 only (difference = $0.79 \pm 0.26\%$, $P = 0.001$), and there was a greater percentage of rule breaks in week 2 than in week 1 in session 2 only (difference = $0.24 \pm 0.07\%$, $P = 0.001$).

Redundancy

Redundancy increased across sessions during the course of the sleep deprivation (Fig. 4). There was a significant main effect of session [$F(13, 1016) = 10.91$, $P < 0.0001$], and comparison

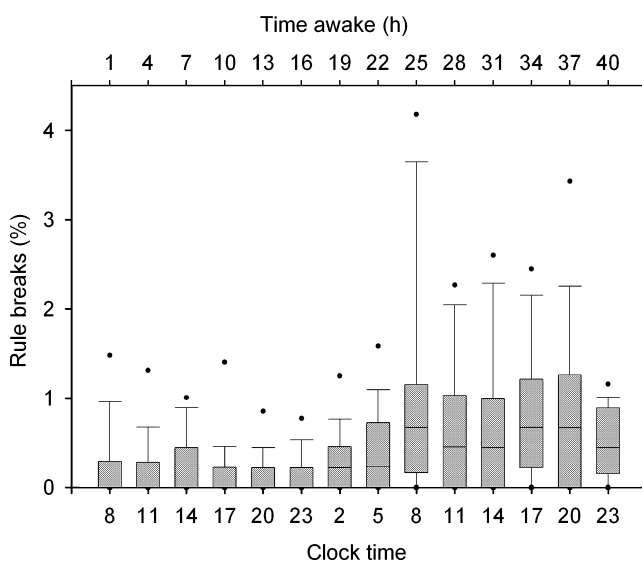


Figure 3. Boxplots of the percentage of rule breaks plotted across the 14 sessions that occurred during sleep deprivation. The middle line shows the median, the outer edges of the boxes indicate the 25th and 75th percentiles, the whiskers indicate the 90th and 10th percentiles and the dots indicate the 5th and 95th percentiles.

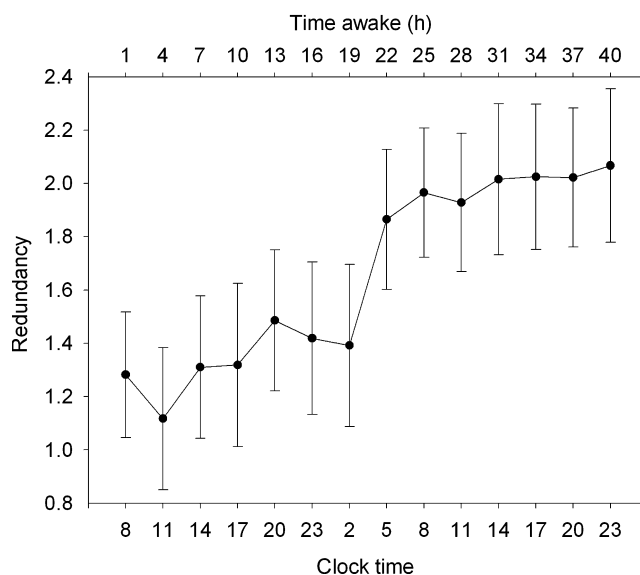


Figure 4. Redundancy, a measure of repetitive usage of response alternatives, increased over the 14 test sessions during sleep deprivation. The ordinate units represent standard deviations from the mean values derived from computer-generated sequences (see Methods). The error bars represent standard errors.

of sessions at analogous times of day revealed a significant 59% increase in redundancy after the night of sleep deprivation (estimated difference = 0.74 ± 0.09 , $P < 0.0001$; means and standard errors, before: 1.25 ± 0.27 , after: 1.99 ± 0.26). Responses generated at the fast tempo were more redundant than responses generated at the slow tempo [$F(1, 1016) = 76.08$, $P < 0.0001$; means and standard errors, fast: 1.91 ± 0.23 , slow: 1.40 ± 0.30].

Null score quotient

The NSQ, a measure of stereotypy in adjacent response pairs, also increased during the course of sleep deprivation (Fig. 5). There was a significant main effect of session [$F(13, 1016) = 3.98$, $P < 0.0001$], and comparison of sessions at analogous times of day revealed a significant 20% increase in NSQ after the night of sleep deprivation (estimated difference = 0.80 ± 0.14 , $P < 0.01$; before: 3.93 ± 0.49 , after: 4.71 ± 0.56). The NSQ was higher for responses generated at the fast tempo than for responses generated at the slow tempo [$F(1, 1016) = 56.38$, $P < 0.0001$; fast: 4.74 ± 0.50 , slow: 4.01 ± 0.56]. There was a speed \times week interaction [$F(1, 1016) = 4.25$, $P = 0.04$], because the NSQ increased across the 2 weeks for the slow speed (slow week 1: 3.68 ± 0.53 , slow week 2: 4.27 ± 0.58 , $P = 0.03$) but did not differ across weeks for the fast speed (fast week 1: 4.75 ± 0.48 , fast week 2: 4.80 ± 0.53 , $P = 0.74$).

Adjacency

The adjacency, a measure of counting tendency, was higher for the fast speed than for the slow speed [$F(1, 767) = 310.3$,

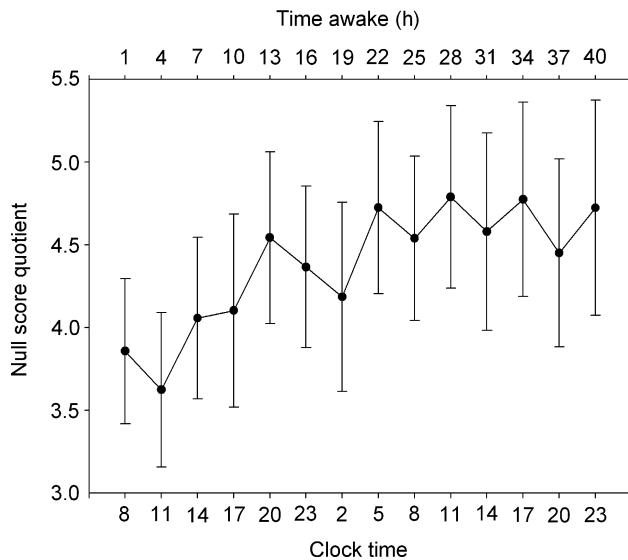


Figure 5. The null score quotient, a measure of stereotypy in response pairs, increased over the 14 test sessions during sleep deprivation. The ordinate units represent standard deviations from the mean values derived from computer-generated sequences (see Methods). The error bars represent standard errors.

$P < 0.0001$; fast: 2.62 ± 0.47 , slow: 0.99 ± 0.46]. Adjacency decreased across weeks [$F(1, 19) = 49.18$, $P < 0.0001$; week 1: 2.64 ± 0.48 , week 2: 0.96 ± 0.45] and also changed across sessions [$F(13, 244) = 1.79$, $P = 0.04$]. As shown in Fig. 6, the latter changes depended on speed and week. On the fast portion of the test, adjacency decreased across sessions in the first week, whereas on the slow portion adjacency did not change across sessions. As illustrated in Fig. 7, the speed \times condition \times week interaction was significant [$F(1, 767) = 16.62$, $P < 0.0001$]. Caffeine tended to lower the adjacency score on the slow portion of the task in the first week only.

DISCUSSION

Sleep deprivation was associated with a larger percentage of rule violations and increased stereotypy of responding in oral RNG. Caffeine did not mitigate these impairments, whereas it did help to prevent drops in the number of numbers generated on the second day of sleep deprivation. Some of the reported beneficial effects of caffeine on cognitive performance may represent the relief of withdrawal symptoms in habitual caffeine users (e.g. James, 1998; James and Gregg, 2004). In the present study, subjects abstained from caffeine for 2 weeks before the study. Thus, the effect observed is unlikely to be attributable to relief of withdrawal. Our findings are consistent with previous evidence that caffeine helps to maintain relatively simple aspects of performance during sleep deprivation (e.g. Beaumont *et al.*, 2001; Bonnet *et al.*, 1995; Lagarde *et al.*, 2000; Lieberman *et al.*, 2002; Lorist *et al.*, 1994; Patat *et al.*, 2000; Reyner and Horne, 2000; Tharion *et al.*, 2003; Wesensten *et al.*, 2002; Wright *et al.*, 1997; Wyatt *et al.*, 2004), while

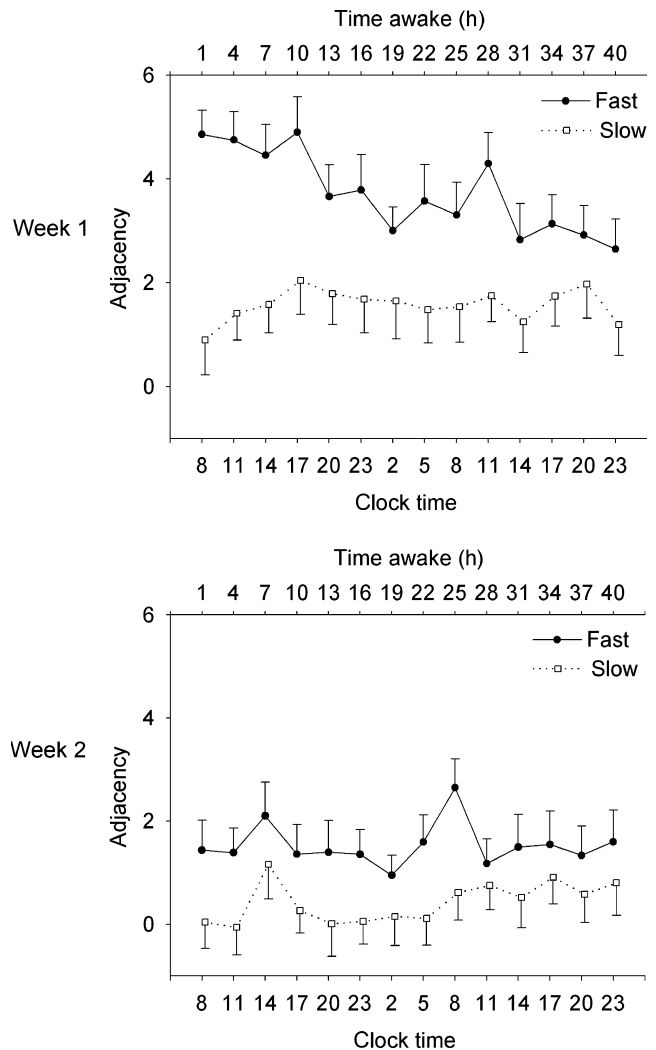


Figure 6. Adjacency, a measure of counting tendency, did not show consistent changes with sleep deprivation. The ordinate units represent standard deviations from the mean values derived from computer-generated sequences (see Methods). The error bars represent standard errors. The session effect depended on speed and week [session \times speed: $F(13, 767) = 2.59$, $P = 0.002$; session \times week: $F(13, 772) = 2.76$, $P < 0.001$]. ANOVAS with the factors session, week, and condition were computed for the fast and slow data separately. For the *fast tempo* (solid symbols), there was a significant main effect of session [$F(13, 245) = 3.45$, $P < 0.0001$] and week [$F(1, 19) = 67.52$, $P < 0.0001$] and a significant session \times week interaction [$F(13, 244) = 3.66$, $P < 0.0001$]. Adjacency decreased across sessions in the first week [upper graph, solid symbols; session main effect: $F(13, 246) = 4.90$, $P < 0.0001$; comparison of sessions 2–4 with sessions 10–12 revealed a significant decrease; estimated difference = 1.24 ± 0.29 , $P < 0.0001$]. In the second week, the adjacency differed among sessions [lower graph, solid symbols; $F(13, 242) = 1.88$, $P = 0.03$] but did not differ between sessions at analogous times of day ($P = 0.37$). For the *slow tempo* (open symbols), the adjacency differed between weeks [$F(1, 19) = 18.94$, $P < 0.001$] and conditions [$F(1, 19) = 6.60$, $P = 0.02$, caffeine: 0.70 ± 0.44 , placebo 1.27 ± 0.56 ; see Fig. 7] but did not change during sleep deprivation [session main effect: $F(13, 243) = 1.38$, $P = 0.17$].

it may not prevent decrements in some complex cognitive tasks (Harrison and Horne, 2000b). Stimulants may help to sustain some, but not necessarily all, complex cognitive abilities during

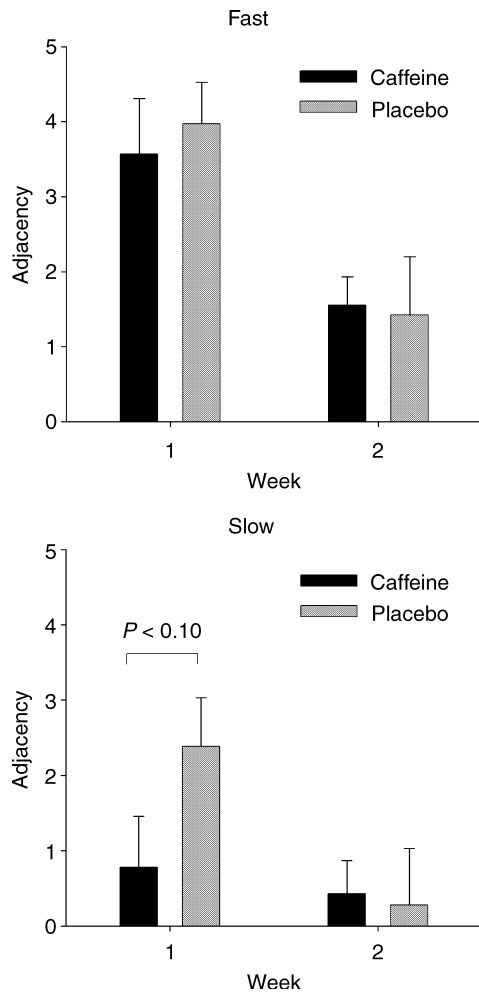


Figure 7. Caffeine was associated with a marginally significant decrease in adjacency scores during the first week on the slow portion of the task. In addition to the three-way interaction of speed \times condition \times week (see text), the four-way ANOVA revealed significant two-way interactions of speed \times week [$F(1, 767) = 28.5, P < 0.0001$] and speed \times condition [$F(1, 767) = 11.54, P < 0.001$]. The ordinate units represent standard deviations from the mean values derived from computer-generated sequences (see Methods). The error bars represent standard errors.

sleep deprivation (Baranski and Pigeau, 1997; Bard *et al.*, 1996; Walsh *et al.*, 2004).

We found that subjects tended to generate fewer numbers than requested and that the negative deviations from the requested sample size became more pronounced with increasing sleep deprivation. Compared with baseline sessions at analogous times of day, subjects in the placebo condition generated on average 8.8 fewer responses during sleep deprivation. These drops in responding may reflect attentional lapses, such as those observed with visual and auditory vigilance tasks present during sleep deprivation or sleep restriction (e.g. Glenville *et al.*, 1978; Van Dongen *et al.*, 2003; Wright *et al.*, 2002). The variability in the placebo condition increased in the later sessions; presumably this variability reflects individual differences in the response to

sleep deprivation (cf. Van Dongen *et al.*, 2003). Caffeine is known to improve reaction times and lapses induced by sleep deprivation (e.g. Patat *et al.*, 2000; Wyatt *et al.*, 2004; Zwyghuizendoorenbos *et al.*, 1990). Similarly, in our study caffeine attenuated the negative deviations from the required sample size and reduced variability. In sessions at analogous times of day, the negative deviation associated with sleep deprivation was 21 times greater in the placebo condition than in the caffeine condition.

Our results support the prediction that sleep deprivation increases stereotypy of responding in oral RNG. Sleep deprivation affected two independent measures of response stereotypy. Compared with baseline sessions at analogous times of day, sleep deprivation was associated with: (1) a 59% increase in redundancy of responding, indicating differential frequency of the usage of response choices regardless of their position in the sequence; and (2) a 20% increase in the NSQ, indicating increased stereotypy of adjacent response pairs. Thus, both zero-order and first-order measures of response stereotypy increased during extended wakefulness. Consistent with these findings, observations of performance on other types of tasks indicate that sleep deprivation (Harrison and Horne, 1997, 1999; Horne, 1988) and sleep restriction (Herscovitch *et al.*, 1980) cause perseveration.

The percentage of rule violations also increased during sleep deprivation. Impairments in the ability to follow task rules (cf. Riccio *et al.*, 2004) and increased perseveration (Harrison and Horne, 1997, 1999; Horne, 1988) may reflect sleep deprivation induced alterations in executive/prefrontal cortical functions. Previous studies suggest that oral RNG assesses executive functions that rely on the frontal and particularly prefrontal cortex (Artiges *et al.*, 2000; Itagaki *et al.*, 1995; Jahanshahi *et al.*, 2000). However, additional studies would be necessary to identify the neural bases for the deficits that we observed. Human RNG may involve multiple overlapping neuroanatomical systems that contribute to different aspects of performance (cf. Daniels *et al.*, 2003), such as keeping track of allowed responses, avoiding response stereotypy and suppressing habitual tendencies such as counting. Sleep deprivation may differentially affect these systems.

Sleep deprivation influenced sample size, rule adherence and zero- and first-order response stereotypy. However, the results did not support the prediction that sleep deprivation would increase counting tendency. A recent study that involved three experiments with random keypressing tasks and one experiment with RNG and random noun generation tasks likewise suggested that sleep deprivation differentially affects various functions that can be assessed with random generation tasks (Heuer *et al.*, 2005). Moreover, the rate of responding can influence the results, and the effects of sleep deprivation may depend on the number of response alternatives and on the type of response solicited (Heuer *et al.*, 2005; Sagaspe *et al.*, 2003). Counting is one of the most highly automated responses associated with oral generation of numbers, but participants easily recognize their tendency to count and try to suppress

counting when attempting to generate random sequences. In our study, the adjacency measure did not show consistent effects across the course of sleep deprivation, and there was evidence for a decrease in counting tendency across weeks. There was also a decrease across sessions in the first week on the fast portion of the task. It remains possible that sleep deprivation increases counting tendency and that practice on the task counteracts this effect. A recent study that involved only two administrations of a RNG task given before and after an experimental night involving either a night of sleep deprivation or normal sleep at home found that counting tendency declined across sessions in the control group but did not change in the sleep deprivation group (Heuer *et al.*, 2005). Our results suggest that caffeine helps to suppress counting tendency before subjects have extensive practice because caffeine was associated with a marginally significant decrease in the adjacency score in the first week for the slow portion of the task.

The present study included only male subjects which limits the generality of the conclusions. Also, because the order of sleep deprivation and the baseline condition (the first day of wakefulness) were not counterbalanced, one cannot rule out the possibility that order effects could have influenced the results. Order effects may influence the outcome of many studies in which tasks are administered repeatedly to investigate the effects of sleep homeostasis or circadian rhythms on cognitive performance. For example, repeated performance of a task may cause boredom and contribute to worsening of performance over time. Interactions between boredom and sleepiness are interesting in their own right (Hayashi *et al.*, 1998), and these interactions are practically relevant (Weinger, 1999). However, boredom alone is unlikely to explain our results. For those variables on which we observed detrimental effects of sleep deprivation, there were generally no changes in performance across the 2 weeks (with the exception of NSQ at the slow speed). If subjects' performance worsened because of boredom, one would expect worse performance in the second week. Yet, there was overall performance improvement across the 2 weeks on both the rule breaks and the adjacency measures, indicating that changes in motivation did not cause a generalized performance decrement. Future studies can address these issues by including women as participants, using counterbalanced designs and collecting subjective ratings of motivation/boredom.

SUMMARY AND CONCLUSIONS

The present results provide evidence that sleep deprivation impairs RNG. Sleep deprivation was associated with drops in the number of responses generated, more violations of task rules and increases in zero- and first-order response stereotypy. RNG is a simple task that may provide a sensitive new method for the investigation of cognitive impairments during sleep deprivation. Caffeine attenuated the deficit in the number of responses generated, but did not mitigate the other impairments in RNG that were associated with sleep deprivation.

Our results support the conclusion that while caffeine helps to sustain simple cognitive processes during sleep deprivation, it may not prevent decrements in more complex aspects of cognitive performance. Further studies are needed to fully characterize the relative efficacy of caffeine and other stimulants in sustaining complex cognitive processes during sleep deprivation.

ACKNOWLEDGEMENTS

This research was supported by Swiss National Science Foundation grants 3100-067060.01 (to HPL) and 3100A0-10567 (to PA), and by a Kirschstein National Research Service Award IF32HL77066-01 from the National Heart, Lung, and Blood Institute (to JMG). We thank J. Buckelmüller and K. Tönz for help with the data collection; T. Jäger for careful transcription of the data; C. Buser and P. Gaccione for expert statistical advice; Dr P. Brugger, Dr R. Dürr, and Dr M. Gottselig for helpful suggestions; and Dr A. Borbély for comments on the manuscript. Statistical analyses were supported in part by the Biostatistics Consulting Service, Center for Clinical Investigation, Brigham and Women's Hospital.

REFERENCES

- Anderson, C. and Horne, J. A. Prefrontal cortex: links between low frequency delta EEG in sleep and neuropsychological performance in healthy, older people. *Psychophysiology*, 2003, 40: 349–357.
- Artiges, E., Salame, P., Recasens, C., Poline, J.-B., Attar-Levy, D., De la Raillere, A., Paillere-Martinot, M. L., Danion, J.-M. and Martinot, J.-L. Working memory control in patients with schizophrenia: a PET study during a random number generation task. *Am. J. Psychiatry*, 2000, 157: 1517–1519.
- Baddeley, A., Emslie, H., Kolodny, J. and Duncan, J. Random generation and the executive control of working memory. *Q. J. Exp. Psychol. Sect. A Hum. Exp. Psychol.*, 1998, 51: 819–852.
- Baranski, J. V. and Pigeau, R. A. Self-monitoring cognitive performance during sleep deprivation: effects of modafinil, d-amphetamine and placebo. *J. Sleep Res.*, 1997, 6: 84–91.
- Bard, E. G., Sotillo, C., Anderson, A. H., Thompson, H. S. and Taylor, M. M. The DCIEM Map Task Corpus: spontaneous dialogue under sleep deprivation and drug treatment. *Speech Commun.*, 1996, 20: 71–84.
- Beaumont, M., Batejat, D., Pierard, C., Coste, O., Doireau, P., Van Beers, P., Chauffard, F., Chassard, D., Enslin, M., Denis, J. B. and Lagarde, D. Slow release caffeine and prolonged (64-h) continuous wakefulness: effects on vigilance and cognitive performance. *J. Sleep Res.*, 2001, 10: 265–276.
- Bonnet, M. H., Gomez, S., Wirth, O. and Arand, D. L. The use of caffeine versus prophylactic naps in sustained performance. *Sleep*, 1995, 18: 97–104.
- Borbély, A. A. and Achermann, P. Sleep homeostasis and models of sleep regulation. In: M. H. Kryger, T. Roth and W. C. Dement (Eds) *Principles and Practice of Sleep Medicine*. Elsevier Saunders, Philadelphia, PA, 2005: 405–417.
- Brown, H. and Prescott, R. *Applied Mixed Models in Medicine*. John Wiley & Sons, Chichester, England, 1999.
- Brugger, P., Monsch, A. U., Salmon, D. P. and Butters, N. Random number generation in dementia of the Alzheimer type: a test of frontal executive functions. *Neuropsychologia*, 1996, 34: 97–103.
- Buchsbaum, M. S., Mendelson, W. B., Duncan, W. C., Coppola, R., Kelsoe, J. and Gillin, J. C. Topographic cortical mapping of EEG

- sleep stages during daytime naps in normal subjects. *Sleep*, 1982, 5: 248–255.
- Cajochen, C., Foy, R. and Dijk, D.-J. Frontal predominance of a relative increase in sleep delta and theta EEG activity after sleep loss in humans. *Sleep Res. Online*, 1999, 2: 65–69.
- Daniels, C., Witt, K., Wolff, S., Jansen, O. and Deuschl, G. Rate dependency of the human cortical network subserving executive functions during generation of random number series – a functional magnetic resonance imaging study. *Neurosci. Lett.*, 2003, 345: 25–28.
- Drummond, S. P. A., Brown, G. G., Stricker, J. L., Buxton, R. B., Wong, E. C. and Gillin, J. C. Sleep deprivation-induced reduction in cortical functional response to serial subtraction. *Neuroreport*, 1999, 10: 3745–3748.
- Drummond, S. P. A., Brown, G. G., Gillin, J. C., Stricker, J. L., Wong, E. C. and Buxton, R. B. Altered brain response to verbal learning following sleep deprivation. *Nature*, 2000, 403: 655–657.
- Drummond, S. P. A., Gillin, J. C. and Brown, G. G. Increased cerebral response during a divided attention task following sleep deprivation. *J. Sleep Res.*, 2001, 10: 85–92.
- Evans, F. J. and Graham, C. Subjective random number generation and attention deployment during acquisition and overlearning of a motor skill. *Bull. Psychon. Soc.*, 1980, 15: 391–394.
- Finelli, L. A., Borbély, A. A. and Achermann, P. Functional topography of the human nonREM sleep electroencephalogram. *Eur. J. Neurosci.*, 2001, 13: 2282–2290.
- Frith, C. D., Friston, K., Liddle, P. F. and Frackowiak, R. S. Willed action and the prefrontal cortex in man: a study with PET. *Proc. R. Soc. Lond. B. Biol. Sci.*, 1991, 244: 241–246.
- Glenville, M., Broughton, R., Wing, A. M. and Wilkinson, R. T. Effects of sleep deprivation on short duration performance measures compared to the Wilkinson auditory vigilance task. *Sleep*, 1978, 1: 169–176.
- Harrison, Y. and Horne, J. A. Sleep deprivation affects speech. *Sleep*, 1997, 20: 871–877.
- Harrison, Y. and Horne, J. A. One night of sleep loss impairs innovative thinking and flexible decision making. *Organ. Behav. Hum. Decis. Process.*, 1999, 78: 128–145.
- Harrison, Y. and Horne, J. A. The impact of sleep deprivation on decision making: a review. *J. Exp. Psychol. Appl.*, 2000a, 6: 236–249.
- Harrison, Y. and Horne, J. A. Sleep loss and temporal memory. *Q. J. Exp. Psychol. A*, 2000b, 53: 271–279.
- Hayashi, M., Minami, S. and Hori, T. Masking effect of motivation on ultradian rhythm. *Percept. Mot. Skills*, 1998, 86: 127–136.
- Herscovitch, J., Stuss, D. and Broughton, R. Changes in cognitive processing following short-term cumulative partial sleep deprivation and recovery oversleeping. *J. Clin. Neuropsychol.*, 1980, 2: 301–319.
- Heuer, H., Kohlisch, O. and Klein, W. The effects of total sleep deprivation on the generation of random sequences of key-presses, numbers and nouns. *Q. J. Exp. Psychol. Sect. A Hum. Exp. Psychol.*, 2005, 58: 275–307.
- Horne, J. A. Sleep loss and divergent thinking ability. *Sleep*, 1988, 11: 528–536.
- Horne, J. A. Human sleep, sleep loss and behavior – implications for the prefrontal cortex and psychiatric-disorder. *Br. J. Psychiatry*, 1993, 162: 413–419.
- Horne, R. L., Evans, F. J. and Orne, M. T. Random number generation, psychopathology, and therapeutic change. *Arch. Gen. Psychiatry*, 1982, 39: 680–683.
- Itagaki, F., Niwa, S. I., Itoh, K. and Momose, T. Random number generation and the frontal cortex. *Int. J. Psychophysiol.*, 1995, 19: 79–80.
- Jahanshahi, M. and Dirnberger, G. The left dorsolateral prefrontal cortex and random generation of responses: studies with transcranial magnetic stimulation. *Neuropsychologia*, 1999, 37: 181–190.
- Jahanshahi, M., Profice, P., Brown, R. G., Ridding, M. C., Dirnberger, G. and Rothwell, J. C. The effects of transcranial magnetic stimulation over the dorsolateral prefrontal cortex on suppression of habitual counting during random number generation. *Brain*, 1998, 121: 1533–1544.
- Jahanshahi, M., Dirnberger, G., Fuller, R. and Frith, C. D. The role of the dorsolateral prefrontal cortex in random number generation: a study with positron emission tomography. *Neuroimage*, 2000, 12: 713–725.
- James, J. E. *Understanding Caffeine: A Biobehavioral Analysis*. Sage Publications, Thousand Oaks, CA, 1997.
- James, J. E. Acute and chronic effects of caffeine on performance, mood, headache, and sleep. *Neuropsychobiology*, 1998, 38: 32–41.
- James, J. E. and Gregg, M. E. Effects of dietary caffeine on mood when rested and sleep restricted. *Hum. Psychopharmacol. Clin. Exp.*, 2004, 19: 333–341.
- Lagarde, D., Batejat, D., Sicard, B., Trocherie, S., Chassard, D., Enslin, M. and Chauffard, F. Slow-release caffeine: a new response to the effects of a limited sleep deprivation. *Sleep*, 2000, 23: 651–661.
- Landolt, H.-P., Rétey, J. V., Tonz, K., Gottselig, J. M., Khatami, R., Buckelmüller, I. and Achermann, P. Caffeine attenuates waking and sleep electroencephalographic markers of sleep homeostasis in humans. *Neuropsychopharmacology*, 2004, 29: 1–7.
- Lieberman, H. R., Tharion, W. J., Shukitt-Hale, B., Speckman, K. L. and Tulley, R. Effects of caffeine, sleep loss, and stress on cognitive performance and mood during US Navy SEAL training. *Psychopharmacology (Berl)*, 2002, 164: 250–261.
- Lorist, M. M., Snel, J. and Kok, A. Influence of caffeine on information processing stages in well rested and fatigued subjects. *Psychopharmacology (Berl)*, 1994, 113: 411–421.
- Matsumoto, M. and Nishimura, T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulation*, 1998, 8: 3–30.
- Patat, A., Rosenzweig, P., Enslin, M., Trocherie, S., Miget, N., Bozon, M. C., Allain, H. and Gandon, J. M. Effects of a new slow release formulation of caffeine on EEG, psychomotor and cognitive functions in sleep-deprived subjects. *Hum. Psychopharmacol. Clin. Exp.*, 2000, 15: 153–170.
- Reyner, L. A. and Horne, J. A. Early morning driver sleepiness: effectiveness of 200 mg caffeine. *Psychophysiology*, 2000, 37: 251–256.
- Riccio, C. A., Wolfe, M. E., Romine, C., Davis, B. and Sullivan, J. R. The Tower of London and neuropsychological assessment of ADHD in adults. *Arch. Clin. Neuropsychol.*, 2004, 19: 661–671.
- Sagaspe, P., Charles, A., Taillard, J., Bioulac, B. and Philip, P. Inhibition and working memory: effect of acute sleep deprivation on a random generation task. *Can. J. Exp. Psychol. Rev. Can. Psychol. Exp.*, 2003, 57: 265–273.
- Spatz, J. and Goldenberg, G. Components of random generation by normal subjects and patients with dysexecutive syndrome. *Brain Cogn.*, 1993, 23: 231–242.
- Tharion, W. J., Shukitt-Hale, B. and Lieberman, H. R. Caffeine effects on marksmanship during high-stress military training with 72 hour sleep deprivation. *Aviat. Space Environ. Med.*, 2003, 74: 309–314.
- Thomas, M., Sing, H., Belenky, G., Holcomb, H., Mayberg, H., Dannals, R., Wagner, H., Thorne, D., Popp, K., Rowland, L., Welsh, A., Balwinski, S. and Redmond, D. Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain activity. *J. Sleep Res.*, 2000, 9: 335–352.
- Towse, J. N. and Neil, D. Analyzing human random generation behavior: a review of methods used and a computer program for describing performance. *Behav. Res. Methods Instr. Comput.*, 1998, 30: 583–591.
- Van Dongen, H. P., Maislin, G., Mullington, J. M. and Dinges, D. F. 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: 117–126.

- Walsh, J. K., Randazzo, A. C., Stone, K. L. and Schweitzer, P. K. Modafinil improves alertness, vigilance, and executive function during simulated night shifts. *Sleep*, 2004, 27: 434–439.
- Weinger, M. B. Vigilance, boredom, and sleepiness. *J. Clin. Monit. Comput.*, 1999, 15: 549–552.
- Werth, E., Achermann, P. and Borbély, A. A. Fronto-occipital EEG power gradients in human sleep. *J. Sleep Res.*, 1997, 6: 102–112.
- Wesensten, N. J., Belenky, G., Kautz, M. A., Thorne, D. R., Reichardt, R. M. and Balkin, T. J. Maintaining alertness and performance during sleep deprivation: modafinil versus caffeine. *Psychopharmacology (Berl)*, 2002, 159: 238–247.
- Wiegersma, S., Vanderscheer, E. and Hijman, R. Subjective ordering, short-term memory, and the frontal lobes. *Neuropsychologia*, 1990, 28: 95–98.
- Wright, K. P., Badia, P., Myers, B. L. and Plenzler, S. C. Combination of bright light and caffeine as a countermeasure for impaired alertness and performance during extended sleep deprivation. *J. Sleep Res.*, 1997, 6: 26–35.
- Wright, K. P., Hull, J. T. and Czeisler, C. A. Relationship between alertness, performance, and body temperature in humans. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, 2002, 283: R1370–R1377.
- Wyatt, J. K., Cajochen, C., Ritz-De Cecco, A., Czeisler, C. A. and Dijk, D. J. Low-dose repeated caffeine administration for circadian-phase-dependent performance degradation during extended wakefulness. *Sleep*, 2004, 27: 374–381.
- Zwyghuizendoorenbos, A., Roehrs, T. A., Lipschutz, L., Timms, V. and Roth, T. Effects of caffeine on alertness. *Psychopharmacology (Berl)*, 1990, 100: 36–39.