University of Groningen

The effect of 40 hours of constant wakefulness on number comparison performance

Steinborn, Michael B.; Bratzke, Daniel; Rolke, Bettina; Gordijn, Margaretha; Beersma, Domien G. M.; Ulrich, Rolf

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

informa
healthcare

THE EFFECT OF 40 HOURS OF CONSTANT WAKEFULNESS ON NUMBER COMPARISON PERFORMANCE

Michael B. Steinborn,[1] Daniel Bratzke,[1] Bettina Rolke,[1] Marijke C. M. Gordijn,[2] Domien G. M. Beersma,[2] and Rolf Ulrich[1]

[1]*Department of Cognitive and Biological Psychology, University of Tübingen, Tübingen, Germany*
[2]*Department of Chronobiology, Rijksuniversiteit Groningen, Groningen, The Netherlands*

We investigated the effects of sleep loss and circadian rhythm on number comparison performance. Magnitude comparison of single-digits is robustly characterized by a distance effect: Close numbers (e.g., 5 versus 6) produce longer reaction times than numbers further apart (e.g., 2 versus 8). This distance effect is assumed to reflect the difficulty of a comparison process based on an analogous representation of general magnitude. Twelve male participants were required to stay awake for 40 h in a quasi-constant-routine protocol. Response speed and accuracy deteriorated between 00:00 and 06:00 h but recovered afterwards during the next day, indicating a circadian rhythm of elementary cognitive function (i.e., attention and speed of mental processing). The symbolic distance effect, however, did not increase during the nighttime, indicating that neither cumulative sleep loss nor the circadian clock prolongs numerical comparison processes. The present findings provide first evidence for a relative insensitivity of symbolic magnitude processing against the temporal variation in energy state. (Author correspondence: michael.steinborn@uni-tuebingen.de)

Keywords Attention; Circadian rhythm; Reaction time; Sleep loss; Symbolic distance effect

INTRODUCTION

Human cognitive efficiency is greatly affected by state variations within the individual. This has been shown in studies examining the effects of total sleep loss and circadian rhythms on human performance (e.g., Carrier & Monk, 2000; Johnson et al., 1992; Miccoli et al., 2008). Several studies have shown that cognitive performance suffers especially

at circadian nadir times, that is, during the night until early in the morning, as shown, for example, by studies on vigilance and event monitoring (e.g., Harrison et al., 2007; Horne et al., 1983), speeded and complex decision-making (e.g., Killgore et al., 2006; Monk & Carrier, 1997), and motor control (e.g., Edwards et al., 2008; Jasper et al., 2009a, 2009b); especially those tasks are considered vulnerable to both rhythmic and homeostatic effects that put strong demands on cognitive control functions (e.g., Harrison & Horne, 2000; Schmidt et al., 2007, for a review). The literature on this subject can be summarized as follows: mental fatigue caused by sleep loss and circadian phase predominantly impairs selective and sustained attention, the speed of mental processing, and working memory (e.g., Lim & Dinges, 2008; Rogers et al., 2003, for a review). These mental functions have been shown to constitute the key processes that enable efficient performance in the workplace and elsewhere.

Cognitive performance during periods of prolonged wakefulness depends on both homeostatic and rhythmic influences (e.g., Dijk et al., 1992). The homeostatic component refers to the effects of accumulating sleep pressure that continuously increases as a function of time awake. The rhythmic component refers to the effects of the circadian rhythm generated by the nucleus suprachiasmaticus (SCN; Duguay & Cermakian, 2009), a neural control center that gives rise to a variety of biochemical, physiological, and mental rhythms (Beersma & Gordijn, 2007; Cajochen et al., 2003). To account for the effects of both homeostatic and rhythmic influences, formal models of prolonged wakefulness typically predict performance from the combined impact of two independent factors: an oscillatory function with a period of about 24 h, accounting for the circadian rhythm, and a monotonous function that accounts for the accumulation of sleep pressure (e.g., Achermann & Borbély, 2003; Åkerstedt et al., 2008a; Borbély, 1982; Daan et al., 1984).

The aim of this study was to examine the effects of homeostatic (sleep loss) and circadian rhythm on elementary numerical abilities. Animals and humans are considered to have a concept of generalized magnitude that can be used to count things, to determine object size or length, or even to estimate the quantity of uncountable collections (e.g., Dehaene et al., 1998). In everyday life, numbers are used in many ways, for example, to tell time, to buy and to sell things, or even to estimate geometric distances. A growing body of research indicates the processes underlying mental comparison constitute an important mental ability that is required in many work-related domains (Butterworth, 2000; Dehaene, 1997). More precisely, it is well established that both organizational (e.g., shiftwork) and environmental (e.g., task constraints, workload, etc.) factors influence cognitive performance. Neglecting those factors can hamper work efficiency, sometimes with disastrous

consequences. For example, this is the case when an operator has to quickly make a decision based on extracted magnitude information from an altimeter display in an airplane. Impairments due to sleepiness and fatigue could result in failures and thus increase the risk of accidents. As of yet, no study has examined the energetic fluctuations in mental comparison capability. Therefore, the purpose of the present study is to examine the fragility or robustness of mental comparison processes under conditions of prolonged wakefulness.

In a prototypical paradigm to study magnitude comparison, individuals are required to make comparative judgements of stimuli that differ from each other in some kind of numerosity, magnitude, or space. When one compares two simultaneously presented items on a particular dimension, such as size or quantity, the more distant the stimuli the easier the comparison. For example, perceptual comparisons of objects usually require more time when the physical difference between the two stimuli is small than when large (Holyoak, 1977). The numerical distance effect is observed even when the magnitude information of stimuli is not perceptually present in amount, size, or length, but is symbolically represented by the stimuli (e.g., Arabic digits) and thus must be retrieved from semantic memory. This was first observed by Moyer and Landauer (1967) when they measured the time for deciding which of two simultaneously presented digits (ranging from 1 to 9) was larger (symbolic distance paradigm). There was a gradual decline in reaction time (RT) as the difference between the digits in a pair decreased: close numbers (e.g., 5 versus 6) produced longer RTs than numbers further apart (e.g., 8 versus 2). Moyer and Landauer suggested that displayed numerals are converted mentally into analogue magnitudes, and a comparison is then made between these magnitudes in much the same way that comparisons are made between physical stimuli, such as loudness or line length.

The idea proposed by Moyer and Landauer has been generalized to other domains in which stimuli are mentally represented in a serial order (e.g., social status), on a timeline (e.g., historic events), or in geometric space (cf. Leth-Steensen & Marley, 2000, for a review). Presently, there is agreement that number comparison involves two essential mental processes. First, external input must be transformed into an internal number representation, and second, this is followed up by a task-relevant decision process. The first step is considered an automatic and resource-independent process, whereas the second step is considered a central process that requires attentional resources. The resource-dependent, decision-making process has been modeled as a noisy integrator that accumulates evidence about the larger of two digits, based on an abstract mental representation of numerical value (e.g., Poltrock, 1989; Sigman & Dehaene, 2005). Hence, if the numerical representation of the two numbers becomes more noisy during the nighttime and/or as a function of time-awake, the

RIGHTSLINK

accumulation of evidence should take longer (and should become less accurate) as a function of comparison difficulty.

To measure the influence of the circadian clock on task performance and to control for potential masking effects, we employed a 40-h quasi-constant-routine (CR) protocol in which the individuals remained under standardized conditions, including constant wakefulness, dim light exposure (<10 lux), isocaloric meals, and isolated from external time cues. The only difference from other CR protocols is that posture was not semi-recumbent. Participants were allowed to move if necessary for the experiment, but physical activity was kept as low as possible. Since many studies have demonstrated homeostatic and circadian rhythm influences on cognitive performance (e.g., Drummond et al., 2006; Killgore et al., 2006; Nilsson et al., 2005; Sagaspe et al., 2006), we first of all expected a general performance decrement as a function of time-awake (the homeostatic component) that should especially be pronounced during the nighttime (the circadian component). Precisely, since task performance depends—besides the specific ability to discriminate between number values—substantial on general cognitive abilities (e.g., speed of cognitive processing) as well as motor processes, we expected to observe an increase of overall RT and a decline of accuracy during the nighttime. If there is any specific effect on the magnitude comparison process, this should be reflected by an increase in the slope of the symbolic distance function. In particular, difficult comparisons (e.g., 5 versus 6) should suffer more from fatigue than easier comparisons (e.g., 3 versus 7). Under the assumption that number comparison is unaffected, neither by the homeostatic component nor by the circadian clock, no difference over time in the numerical distance effect would be expected.

METHODS AND MATERIALS

Participants

Twelve male volunteers (mean ± SD: age 23.2 ± 3.7 yrs; body weight 78.6 ± 8.6 kg; body mass index [BMI] 22.1 ± 1.6) took part in the study. All were right-handed and had normal or corrected-to-normal vision. They were recruited via advertisement by the Department of Chronobiology of the University of Groningen and were paid for their participation. Only male participants were included, so as to avoid masking effects due to the female menstrual cycle. The selection procedure was based on the following inclusion criteria: (1) no shiftwork, (2) no intercontinental flights the previous 3 months, (3) no regular drug intake and no smoking, (4) no current medical, psychiatric, and sleep disorders, and (5) good health condition. All participants were intermediate chronotypes, as assessed by the Munich Chronotype Questionnaire

(Roenneberg et al., 2003). The protocol was approved by the Medical Ethical Committee of the University Medical Center Groningen, and participants signed an informed consent. It also met international ethical standards for human chronobiology research established by the Journal (Portaluppi et al., 2008).

Protocol

A 40-h quasi-constant-routine (CR) protocol was conducted, starting at waking up at 07:00 h and ending at 23:00 h the next day. Participants slept the night before and after the CR in the lab. During the CR, participants worked on several tasks, including the here-reported digit comparison task and psychomotor vigilance task (PVT) every 3 h. In each of the 13 sessions, the PVT was administered before the digit comparison task, with a rest break between the two tasks. To minimize masking effects due to spontaneous alerting and/or motor activity, participants sat in comfortable chairs in an upright position. They were not allowed to move around, and they left the room only for regular room changes and toilet requirements. Light intensity was kept dim, maximal 10 lux, and participants were isolated from time information or external time cues. To control for metabolic influences due to food intake, isocaloric meals of ~100 kcal/portion and mineral water were provided hourly; caffeinated drinks were not allowed.

Salivary melatonin samples were collected hourly. Saliva samples were assayed for melatonin using a direct double-antibody radioimmunoassay, validated by gas chromatography-mass spectroscopy with an analytic least-detectable concentration of 0.15 pg/mL and in a functional least-detectable concentration of 0.65 pg/mL (Weber et al., 1997). To normalize individual melatonin levels a 3-harmonic fit was calculated to each individual melatonin curve. DLMO (dim light melatonin onset) was defined as the time when the raw data curve crossed the 25% level of the fitted maximum (by interpolating the last value before and the first value after crossing the 25% criterion).

Self-Ratings

Visual Analog Scales (VAS)

Subjective alertness/sleepiness, motivation, and energy were assessed via visual analog scales (VAS), which are typically used to measure subjective states (e.g., moods, feelings, etc.) during CR protocols (Akerstedt & Gillberg, 1990). These measures were assessed hourly, resulting in 40 measurement points available for statistical analysis. Participants had to denote their subjective state, e.g., their self-rated degree of sleepiness,

motivation, and energy at the respective moment, by marking a position on a continuous line (10 cm) between two endpoints. The left endpoint marked the lowest level and the right endpoint marked the highest level of the respective scale. To ease comparison of the three self-rating measures across the sessions, the polarity of the sleepiness scale was inversed (indicating the participants' alertness). To account for scaling effects, normalized values (proportional deviation from individual mean across sessions) of all three VAS variables were used to compose figures and conduct statistical analyses.

Assessment of Cognitive Performance

Psychomotor Vigilance Task (PVT)

As a baseline condition, we used a 10-min computerized version of the PVT, which is a simple RT task assumed to measure the participants' basic alertness level. The PVT has become the standard laboratory tool for the assessment of sustained performance in a variety of experimental conditions. Participants were seated ∼60 cm in front of a grey (5.0 cd/m^2) computer screen. During the task, they responded to a visual imperative stimulus (IS) consisting of a white square (100 cd/m^2, $1.10° \times 1.10°$ of visual angle) by pressing the space-bar with their right index finger. The response terminated the IS. If no response occurred within 2 s, the IS was terminated. A variable inter-stimulus interval (3, 4, 5, 6, or 7 s) separated subsequent trials (Langner et al., 2010). Participants performed 120 experimental trials in each of the 13 experimental sessions scheduled 3 h apart from one another.

Symbolic Distance Paradigm

We employed a single-digit selection version of the symbolic distance paradigm in which participants were required to choose the larger of two simultaneously presented single-digit numbers. From the 72 possible combinations, 16 digit-pairs were selected (9-1, 1-9, 9-2, 2-9, 8-2, 2-8, 8-3, 3-8, 7-3, 3-7, 7-4, 4-7, 6-4, 4-6, 6-5, and 5-6). By means of this selection, we controlled for a frequency bias, that is, the fact there are more combinations possible for the difficult comparisons than for easy ones. Item difficulty is reflected in the numerical distance between a pair of digits, ranging from 1 (e.g., 6-5) to 8 (e.g., 1-9). In addition, we ensured the responses were counterbalanced so that there was an equal number of right versus left responses. Participants were instructed to quickly and accurately decide the larger of two simultaneously presented digits; they had to respond with either the left shift-key (left index finger when the

leftward presented digit was larger) or the right shift-key (right index finger when the rightward presented digit was larger).

With the participants seated $\sim$60 cm in front of a blue (7.1 cd/m^2) computer screen, the experiment started with the presentation of a fixation line ("$_\ _\ _\ _\ _$"; 2.3° angle of vision) in the center of the screen, followed by a blank interval of 500 ms duration, after which the IS (for each digit, white color, 100 cd/m^2, 1.14° × 0.86° of visual angle) occurred. The fixation line was consistently present throughout the experimental session but was replaced by the IS. The IS was presented in the center of the screen and was terminated either by the participant's response or when the response interval expired after 2000 ms. A response-stimulus interval of 500 ms separated subsequent trials. Participants were instructed to respond quickly and accurately to the IS. Feedback was given if an erroneous response had occurred. In the case of an erroneous response, the Dutch word "fout" (wrong) was presented for 300 ms. Items were presented randomly, with each of the 16 digit-pairs occurring 30 times during the course of a single session. Thus, the participants performed 480 trials/session, lasting about 10 min.

Apparatus

Both the single-digit selection task and the PVT were run on a Fujitsu-Siemens Notebook with 19-inch color TFT-display; it was programmed using the Experimental Runtime System (ERTS) software package.

Procedure

After reading the written information of the study and signing the informed consent, participants were invited to the lab 1 week before the beginning of the CR protocol. They were given their final instructions and received a small, rugged, actigraphy-based data logger (actiwatch) that records a digitally integrated measure of gross motor activity. To control for sleep-wake habits, participants were required to keep a sleep diary. To familiarize participants with the tasks and to reduce initial practice effects, participants performed a practice session 1 day before the CR protocol. During the CR protocol, they performed 13 experimental sessions, one every 3 h. Each session began with the administration of questionnaires and melatonin measurement, followed by the PVT and symbolic distance task.

Data Analysis

For the PVT, and for each condition of the symbolic distance task, RTs <100 ms or >2.5 SD from the individual mean were considered

outliers and corresponding trials were discarded from the analysis (~3%). For the PVT, a within-subject analysis of variance (ANOVA) was performed with the factor session (13 levels) and with RT as dependent measure. For the symbolic distance task, the ANOVA contained the factors session (13 levels) and comparison difficulty (8 levels), and RT and error percentage as dependent measures. For self-ratings, the ANOVA contained the factors session (40 levels) and normalized rating scores as dependent measure. p values were adjusted for violations of the assumption of sphericity using the Greenhouse-Geisser correction. For both the PVT and the symbolic distance task, the standard error of mean was computed according to Cousineau (2007), removing individual differences by adjusting the individual means $\bar{X}p$ to the group means $\bar{X}$ for each participant and each session ($Y = X - \bar{X}p + \bar{X}$). An F-tested harmonic regression analysis was used to fit linear combinations of sine and cosine functions to the data (step by step from a 24 h component to higher harmonics until subsequent harmonics no longer significantly contribute to the explained variance). Thus, this additional analysis allowed the assessment of both peak phase and amplitude (Hut, 2007). In order to evaluate the phase relationship between melatonin secretion, self-ratings, and task performance, we calculated cross-correlations between these variables. For these analyses, both melatonin and self-report data were binned into 3 h intervals, like the performance data.

RESULTS

Figure 1 depicts the salivary melatonin secretion, subjective measures of alertness/sleepiness, motivation, and energy as assessed via the VAS (Panel B) and the numerical comparison performance (Panel A) as a function of time-of-day. Figure 2 depicts the performance on PVT (Panel B) and the symbolic comparison task (Panel A) as a function of time-of-day.

Melatonin and Self-Ratings

The DLMO was employed as a circadian phase marker. On average, the DLMO occurred at 20:59 h. The homogeneity of the sample with regard to chronotype is reflected in the relatively small variance across subjects in the DLMO (SD = 46 min). A within-subject ANOVA (factor: session) revealed the self-reported levels of subjective alertness/sleepiness ($F(1, 11) = 15.1$, $\eta^2 = .58$, $p < .001$), motivation ($F(39, 429) = 7.2$, $\eta^2 = .39$, $p < .001$), and energy ($F(39, 429) = 6.7$, $\eta^2 = .38$, $p < .001$) varied considerably across the total session. Subjective states were best at 08:00 h the first day and worst at 06:00 h the second day (contrast for alertness: $F(1, 11) = 53.6$, $\eta^2 = .83$, $p < .001$; motivation: $F(1, 11) = 58.7$, $\eta^2 = .84$,
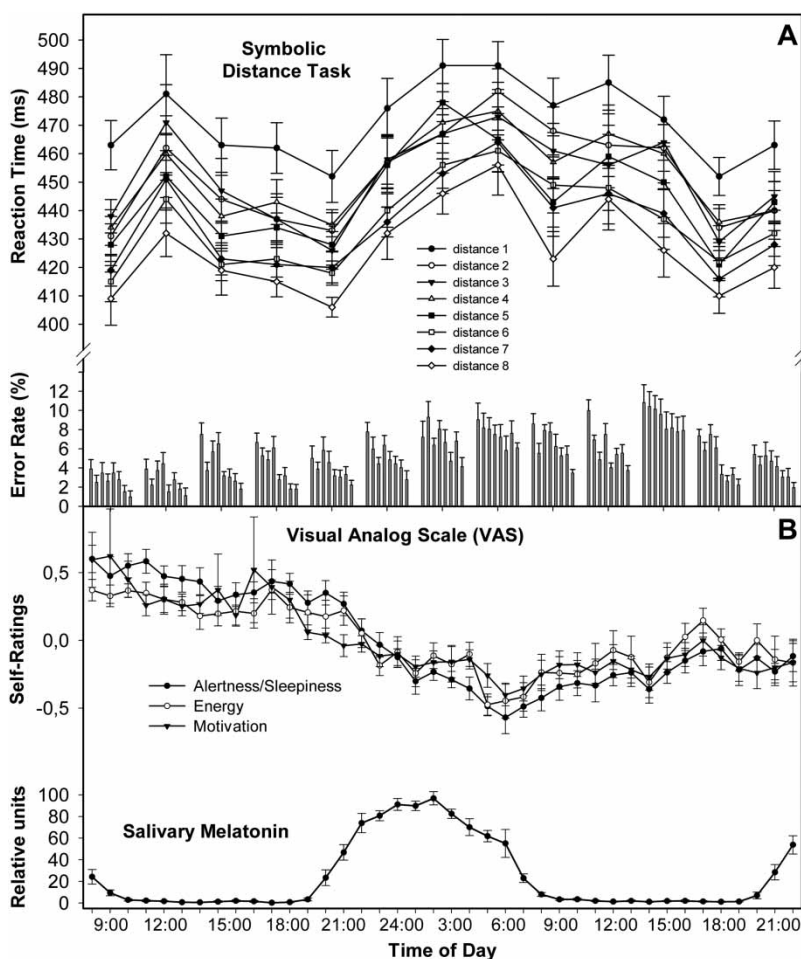
FIGURE 1 Numerical comparison performance, self-ratings, and melatonin secretion as a function of time-of-day. Panel A displays performance speed (mean reaction time) and accuracy (error percentage; from left to right represents from close to far distances) in the symbolic distance task. Panel B displays self-rated energy, motivation, alertness/sleepiness (self-assessed via VAS), and salivary melatonin secretion (relative units, % of the maximum pg/mL), measured hourly. The error bar indicates ± SE.

$p < .001$; energy: $F(1, 11) = 27.0$, $\eta^2 = .71$, $p < .001$). From 06:00 h the second day, subjective states then improved towards 17:00 h (contrast for alertness: $F(1, 11) = 28.6$, $\eta^2 = .72$, $p < .001$; motivation: $F(1, 11) = 13.7$, $\eta^2 = .56$, $p < .01$; energy: $F(1, 11) = 10.7$, $\eta^2 = .49$, $p < .01$) but did not reach the initial level (comparison of 08:00 h [the first day] against 17:00 h [the second day], contrast for alertness: $F(1, 11) = 16.1$, $\eta^2 = .60$, $p < .01$; motivation: $F(1, 11) = 12.6$, $\eta^2 = .53$, $p < .01$; energy: $F(1, 11) = 3.9$, $\eta^2 = .26$, $p < 0.07$). Further, the pattern of self-ratings over time revealed a close phase relationship with the rhythm of melatonin (see
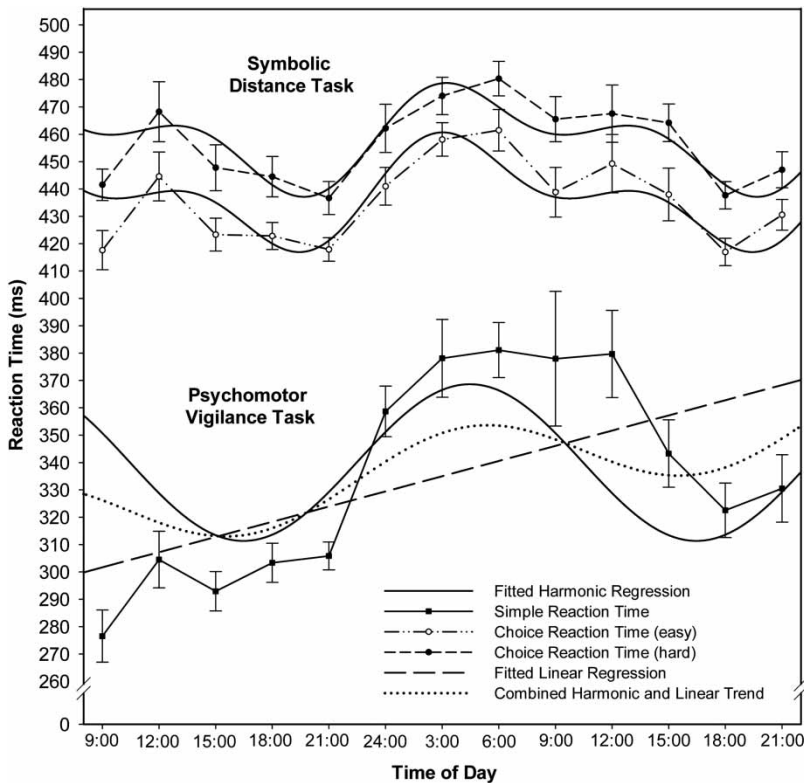
FIGURE 2   Performance of the psychomotor vigilance test (PVT) and symbolic comparison task (aggregated into two levels: easy versus difficult) as a function of time-of-day. The error bar indicates ± SE. The solid lines (sine curves) show the best fit of the experimental data obtained from the harmonic and linear regression analyses. The dotted line shows the combined harmonic and linear trend for the PVT data.

Figure 1). Cross-correlational analyses revealed the highest phase relationship between melatonin and self-ratings at a lag of one session (alertness: $r = -.53$; motivation: $r = -.50$; energy: $r = -.57$).

Psychomotor Vigilance Task

The within-subject ANOVA revealed a significant effect of the factor session on RT ($F(12, 132) = 8.4$, $\eta^2 = .43$, $p < .001$). The fastest responses on the PVT occurred at 09:00 h (RT = 276 ms) the first day and the slowest ones occurred at 06:00 h (RT = 380 ms) the second day ($F(1, 11) = 36.5$, $\eta^2 = .78$, $p < .001$). From 06:00 h the second day, RT performance improved up to a maximum towards 18:00 h (RT = 322 ms) ($F(1, 11) = 12.5$, $\eta^2 = .53$, $p < 0.01$), albeit it did not reach its initial level of the first testing session (comparison of 09:00 h [first day] against 18:00 h [second day], $F(1, 11) = 33.8$, $\eta^2 = .75$, $p < .001$). The harmonic

regression analysis (Hut, 2007) revealed a significant sinusoidal rhythm ($F = 10.7$; $R^2 = .12$; $p < .01$) with an amplitude (one-half the peak-to-trough 24 h variation) of 29 ms (8.4% of mean level) and a fitted peak phase at 16:30 h. Further, a linear regression analysis revealed a significant homeostatic effect on psychomotor vigilance performance ($F = 26.4$; $R^2 = .15$; $p < .001$), with an intercept of 301 ms and a linear regression slope of 2 ms/testing session. Cross-correlational analyses revealed the highest phase relationship between salivary melatonin level and RT at a lag of one session ($r = .66$), and between RT and subjective state at a lag of zero session (alertness: $r = -.91$; motivation: $r = -.90$; energy: $r = -.92$).

Symbolic Distance Task

For the symbolic distance task, ANOVA revealed a significant numerical distance effect on RT ($F(7, 77) = 24.0$, $\eta^2 = .69$, $p < .001$) and error percentage ($F(7, 77) = 13.7$, $\eta^2 = .56$, $p < .001$). The larger the numerical distance between the two digits, the shorter the RT and the lower the error percentage. In addition, there was a main effect of session on overall RT ($F(12, 132) = 4.0$, partial $\eta^2 = .27$, $p < .01$) and on error percentage ($F(12, 132) = 3.1$, partial $\eta^2 = .22$, $p < .05$). The initially fast responses at 09:00 h (RT = 430 ms) the first day became slowest towards 06:00 h (RT = 471 ms) the second day ($F(1, 11) = 19.1$, $\eta^2 = .63$, $p < 0.001$). From 06:00 h the second day, performance then improved towards 18:00 h (RT = 427 ms) in the late afternoon ($F(1, 11) = 8.6$, $\eta^2 = .44$, $p < .01$), indicating a predominant influence of circadian phase on overall RT performance. Error rates were lowest at 09:00 h (2.6%) the first day but continuously increased up to a maximum at 15:00 h (6.5%) the second day ($F(1, 11) = 15.3$, $\eta^2 = .58$, $p < .01$). Most importantly, there was no significant interaction between the factors of session and comparison difficulty, neither for RT nor for error percentage ($F < 1.3$), indicating circadian phase did not affect the numerical distance effect. It should be noted that the results did not change when the ANOVA was performed with only seven distance levels, that is, when the easiest distance condition (i.e., 8, 1 versus 9 and 9 versus 1, respectively) was excluded. This was checked for all statistical procedures to control for a possible end-effect (Leth-Steensen & Marley, 2000), which refers to the finding that responses are usually fastest and most accurate for comparisons at the "ends" of the range of stimuli being compared. Some researchers have argued that "9 versus 1" (and vice versa) can be decided without a comparison process (when the participants look only for "9" and then respond in this direction); others, however, have shown that "9 versus 4" takes longer than "9 versus 1," which is taken as argument that a comparison process must be involved.

To examine the possibility that the influence of circadian phase on numerical comparison is masked by a lack of statistical power, we performed an additional analysis with aggregated data. We categorized the 13 levels of the factor session into three levels: daytime 1 (an aggregate of sessions 1–4 corresponding to the clock times of 09:00, 12:00, 15:00, and 18:00 h), nighttime (an aggregate of sessions 6–8 corresponding to the clock times of 00:00, 03:00, and 06:00 h), and daytime 2 (an aggregate of sessions 9–12 corresponding to the clock times of 09:00, 12:00, 15:00, and 18:00 h the second day). Moreover, we aggregated the eight difficulty levels into *easy* (far distance of 8, 7, 6, 5) versus *difficult* (close distance of 4, 3, 2, 1). A within-subject ANOVA revealed a significant session (daytime versus nighttime) × distance (easy versus difficult) interaction on RT ($F(1, 11) = 7.0$, partial $\eta^2 = .39$, $p < .05$), but only a marginally significant interaction effect on error rate ($F(1, 11) = 2.9$, $\eta^2 = .22$, $p < .10$). RTs were 427 ms (easy) and 451 ms (difficult) for daytime 1, 454 ms (easy) and 472 ms (difficult) for nighttime, and 436 ms (easy) and 459 ms (difficult) for daytime 2. Error rate was 2.1% (easy) and 4.4% (difficult) for daytime 1, 4.4% (easy) and 6.3% (difficult) for nighttime, and 4.0% (easy) and 6.7% (difficult) for daytime 2 (Figure 3). Surprisingly, however, the interaction effect was opposite to our expectation, for both RT and error rate, indicating that circadian phase caused a decrease instead of the expected increase of symbolic distance effect. It should be emphasized that it was necessary to aggregate daytimes 1 and 2 (as daytime, contrasted to nighttime) to obtain sufficient statistical power to reveal an interaction effect. However, since there was a significant interaction in the opposite direction, the additional analysis ruled out the possibility of a hidden effect of circadian phase on comparison difficulty (see Figure 3).

For the linear and harmonic regression analyses, we used the aggregated levels easy comparison (distance 8, 7, 6, 5) versus difficult comparison (distance 4, 3, 2, 1). The harmonic regression analysis revealed a significant sinusoidal rhythm for both the easy ($F = 8.8$; $R^2 = .18$; $p < .01$) and difficult task condition ($F = 7.3$; $R^2 = .16$; $p < .01$). For the easy condition, the amplitude was 22 ms (5% of mean level) and the fitted peak phase was 19:20 h. For the difficult condition, the amplitude was 21 ms (4.3% of mean level) and the fitted peak phase was 19:40 h. However, neither the RT difference between the easy and the difficult condition, nor the slope of the symbolic distance function revealed circadian rhythmicity. Linear regression analysis revealed no significant homeostatic effect on the symbolic distance task, neither for the easy ($F < 1.6$; $R^2 = .00$) nor difficult condition ($F < .6$; $R^2 = .00$), or any other index of numerical comparison performance. Cross-correlational analysis revealed the highest phase relationship between melatonin and RT task performance at a lag of one session (easy: $r = .73$; difficult: $r = .72$), a perfect correspondence between the easy and the difficult condition of
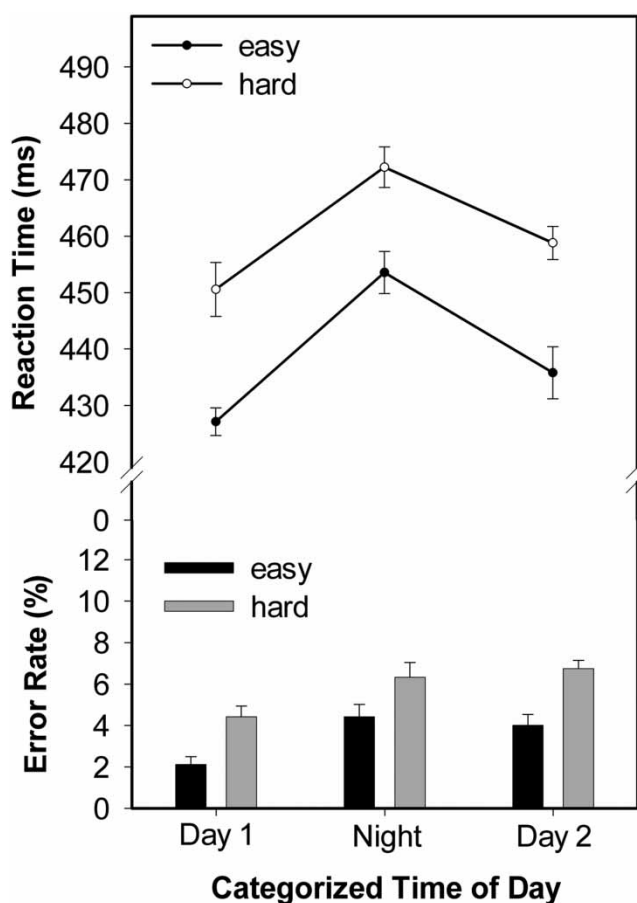
FIGURE 3 Reaction time and error percentage as a function of categorized numerical distance (easy versus difficult) and category of session (daytime 1, night, daytime 2).

the symbolic distance task (no lag, $r = .98$), and between RT and self-ratings at a lag of zero session (alertness: $r = -.59$; motivation: $r = -.58$; energy: $r = -.65$).

DISCUSSION

In the present study, we examined the influence of homeostatic (sleep loss) and rhythmic (circadian clock) influences on symbolic magnitude processing across 40 h of constant wakefulness. To this end, we examined performance on a single-digit comparison paradigm across 13 subsequent sessions, including measures of alertness, self-reports, and melatonin secretion. Harmonic regression analyses revealed a circadian modulation in both simple RT performance (i.e., PVT) and choice RT performance

(i.e., symbolic distance task), with fastest responses observed during the daytime and slowest responses observed during the nighttime. Moreover, there was a close phase relationship between decrements in RT performance, self-report measures of sleepiness (Akerstedt et al., 2008b), and onset of melatonin secretion, which is considered the most important physiological marker of the circadian phase (e.g., Beersma & Gordijn, 2007; Cajochen et al., 2003). The fact that the performance impairments during the nighttime, that is, from 00:00 h to 6:00 h, recovered afterwards during the next day supports the view that overall performance, including perceptual, cognitive, and motor processes, is essentially modulated by the rhythmic activity of the circadian clock and not only by the cumulative effects of sleep loss (Carrier & Monk, 2000; Rogers et al., 2003).

Beside this global performance variation, which is typically observed in choice-RT tasks (e.g., Monk & Carrier, 1997), we asked whether either the circadian phase or the homeostatic component or both also affect the numerical comparison process. If numerical comparison is sensitive to state variations due to the circadian phase and/or the homeostatic component, this should be indicated by a corresponding variation of the numerical distance effect. As displayed in Figure 1A, the performance curves for the eight difficulty conditions were quite parallel across the 13 sessions. Statistically, they did not exhibit any interaction effect on RT that would indicate that difficult comparisons suffered more than easy comparisons at night. Yet, additional analyses with aggregated data revealed that the numerical distance effect even decreased slightly during the nighttime (see Figure 3A). Therefore, since the interaction effect was in the opposite direction, the present results cannot be attributed to a lack of statistical power. Thus, the empirical data suggest the symbolic distance effect is relatively insensitive to both the effects of circadian phase and accumulating sleep pressure.

This finding appears to be counterintuitive and contrasts with other studies that reported a stronger performance decrement during the nighttime with higher task difficulty. For example, Bratzke et al. (2007) recently observed an interaction of task difficulty with circadian phase in a 28-h CR protocol using a dual-task paradigm. Responses slowed down more during the nighttime when the two tasks were presented in close succession compared to when the two tasks were temporally presented more remote from each other. A similar interaction between task difficulty and circadian phase was observed during 40 h of constant wakefulness using a task-switching paradigm (e.g., Bratzke et al., 2009). Performance suffered more when participants had to switch between two tasks compared to when the same task was repeated across subsequent trials. Thus, a theoretically driven explanation of the present results would be that mental comparison of magnitudes is relatively robust against variations in the energetic state of humans.

As mentioned in the Introduction, mental comparison is considered to involve two essential steps: an automatic process of transforming external input into an abstract magnitude representation, and a rather attention-demanding process of coming to a decision. We hypothesized that if the process of accumulating evidence about digit parity becomes noisier during the nighttime, this should result in an increased slope of the symbolic distance effect. The fact that sleepiness had no effect on the symbolic distance function could be interpreted that mental comparison is one of the rather robust mental processes that are relatively insensitive against energetic-state fluctuations. Thus, although the process of accumulating evidence towards the correct response is considered to require attentional resources, part of the decision process may actually be processed automatically–as recently suggested (e.g., Ansari et al., 2006; Oriet et al., 2005).

Although we believe our study provides important new data regarding the effects of sleep loss and circadian rhythm on cognitive processing, several possible caveats must be considered. For example, it could be argued the nighttime decrease of the symbolic distance effect is somehow due to changes in the way individuals perform the task. For example, Dehaene (1997, p. 87) discussed the possibility that, beside a deliberate and attention-demanding way of information processing, there might be a second way of coming to a decision that he labeled a heuristic mode. Once a digit is presented, he argued, individuals quickly have some intuition about the correct response, albeit not with the same degree of confidence. Thus, in situations of limited capacity, as may be induced by constant wakefulness, participants might have shifted from deliberate processing towards a rather less effortful strategy. Sanders (1998, pp. 44–47) and Wilkinson (1990) argued that under conditions of sleep deprivation, participants in a choice-RT situation may not wait until they have accumulated full evidence about the correct response, that is, until they are definitely confident, but may respond earlier on the basis of partial evidence, that is, when they feel somewhat confident. This strategy-shift assumption has been supported by subsequent work by Murphy et al. (2006). In the present study, the continuous decline in overall accuracy in the symbolic distance task is also in line with this explanation (see Figure 1A). Therefore, a tendency towards incomplete information processing during states of fatigue could have prevented the symbolic distance function from increasing during the nighttime and/or as a function of extended wakefulness.

Further, we cannot exclude the possibility the observed performance effects are confounded by practice and by effects of familiarity with the task upon repeated test administration. Specifically, practice effects might have masked possible variations in number comparison as indicated by the symbolic distance effect. Some studies, for example, have not found an interaction between circadian phase and cognitive control and have

RIGHTSLINK

attributed this to repeated-testing effects. For example, Sagaspe et al. (2006) examined the effect of 36 h of constant wakefulness on color-word interference in the Stroop task. In this paradigm, individuals are required to make choice responses to the color of a presented word, whereas the meaning of the word, itself, has to be ignored. Responses are typically fast when color and word meaning are congruent (e.g., BLUE is presented in blue color) compared to when they are incongruent (e.g., RED is presented in blue color). Because there was no interaction between task difficulty and circadian phase in their study, the authors suggested that repeated testing of tasks upon which a target must be focused in the presence of distractors (as is the case in the Stroop task) makes individuals develops a skill of suppressing the distractor, which then masks the true effects of prolonged wakefulness on performance (see Sagaspe et al., 2006, p. 82, for discussion).

We think that a practice-related explanation of the present results is unlikely because several studies have shown performance of the symbolic distance task is remarkably robust against practice effects (e.g., Poltrock, 1989). Even extended practice of the symbolic distance task only slightly decreases overall RT, but it does not change the symbolic distance effect (Link, 1990). Dehaene (1997, p. 74) reappraised these observations in a study with much more practice, in which he extensively trained participants for several days with the symbolic distance paradigm. He was surprised that there was virtually no effect of practice on the distance function, though such an invariance of practice has also been noted for memory search processes (Kristofferson, 1972). The fact that the distance effect decreased from daytime 1 towards nighttime but then increased again during daytime 2 (see Figure 3) may be taken as further evidence against the possibility that practice effects confounded the present results. It could be argued that processes such as number comparison, which are often used in everyday life, become virtually invariant against practice and fatigue (Sternberg, 1998, p. 760).

In conclusion, our results provide evidence that, beyond overall RT performance, numerical comparison is only little affected by sleep loss and circadian rhythms. Whereas several studies observed a pronounced performance decrement with increasing task difficulty, we found the symbolic distance effect remained relatively stable (i.e., even decreased) during 40 h of constant wakefulness. Of course, it would be premature to finally conclude that numerical comparison processes are neither affected by circadian phase nor by cumulative sleep pressure, since there was a circadian rhythm in the performance of the symbolic distance task, albeit it did not interact with comparison difficulty. Further research is needed to examine the effect of extended wakefulness with other variations of the symbolic distance paradigm, or with other paradigms that essentially involve mental comparison, respectively. For example, one might

consider taking concurrent measures of other magnitude judgements, such as perception of length or intensity, to examine the generality of the present findings. One might also consider other approaches to study circadian rhythms on magnitude comparison performance. For example, a forced-desynchrony protocol would make it possible to disentangle the separate contributions of sleep-wake history and circadian phase on comparison performance, which is not possible with CR protocols, since it realizes multiple internal phase relationships between sleep-wake cycles and circadian rhythms (cf. Johnson et al., 1992). Nevertheless, the present study is the first on this subject, and the empirical data so far provide evidence for a relative insensitivity of numerical information processing under conditions of mental fatigue and restricted energy availability.

ACKNOWLEDGMENTS

REFERENCES

Achermann P, Borbély AA. (2003). Mathematical models of sleep regulation. *Front. Biosci.* 8:683–693.

Åkerstedt T, Gillberg M. (1990). Subjective and objective sleepiness in the active individual. *Int. J. Neurosci.* 52:29–37.

Åkerstedt T, Ingre M, Kecklund G, Folkard S, Axelsson J. (2008a). Accounting for partial sleep deprivation and cumulative sleepiness in the Three-Process Model of alertness regulation. *Chronobiol. Int.* 25:309–319.

Åkerstedt T, Kecklund G, Axelsson J. (2008b). Effects of context on sleepiness self-ratings during repeated partial sleep deprivation. *Chronobiol. Int.* 25:271–278.

Ansari D, Dhital B, Siong SC. (2006). Parametric effects of numerical distance on the intraparietal sulcus during passive viewing of rapid numerosity changes. *Brain Res.* 1067:181–188.

Beersma DGM, Gordijn MCM. (2007). Circadian control of the sleep-wake cycle. *Physiol. Behav.* 90:190–195.

Borbély AA. (1982). A two process model of sleep regulation. *Hum. Neurobiol.* 1:195–204.

Bratzke D, Rolke B, Steinborn MB, Ulrich R. (2009). The effect of 40 hours constant wakefulness on task switching efficiency. *J. Sleep Res.* 18:167–172.

Bratzke D, Rolke B, Ulrich R, Peters M. (2007). Central slowing during the night. *Psychol. Sci.* 18:456–461.

Butterworth B. (2000). *The mathematical brain*. London: McMillian.

Cajochen C, Kräuchi K, Wirz-Justice A. (2003). Role of melatonin in the regulation of human circadian rhythms and sleep. *J. Neuroendocrinol.* 15:432–437.

Carrier J, Monk TH. (2000). Circadian rhythms of performance: new trends. *Chronobiol. Int.* 17:719–732.

Cousineau D. (2007). Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Tutorials Quant. Methods Psychol.* 1:42–45.

Daan S, Beersma DGM, Borbely AA. (1984). Timing of human sleep: recovery process gated by a circadian pacemaker. *Am. J. Physiol.* 246:161–183.

Dehaene S. (1997). *The number sense. How the mind creates mathematics*. Oxford, UK: Oxford University Press.

Dehaene S, Dehaene-Lambertz G, Cohen L. (1998). Abstract representations of numbers in the animal and human brain. *Trends Neurosci.* 21:355–361.

Dijk DJ, Duffy JF, Czeisler CA. (1992). Circadian and sleep/wake dependent aspects of subjective alertness and cognitive performance. *J. Sleep Res.* 1:112–117.

Drummond SP, Paulus MP, Tapert SF. (2006). Effects of two nights sleep deprivation and two nights recovery sleep on response inhibition. *J. Sleep Res.* 15:261–265.

Duguay D, Cermakian N. (2009). The crosstalk between physiology and circadian clock proteins. *Chronobiol. Int.* 26:1479–1513.

Edwards B, Waterhouse J, Reilly T. (2008). Circadian rhythms and their association with body temperature and time awake when performance a simple task with the dominant and non-dominant hand. *Chronobiol. Int.* 26:115–132.

Harrison Y, Horne JA. (2000). The impact of sleep deprivation on decision making: a review. *J. Exp. Psychol. Appl.* 6:236–249.

Harrison Y, Jones K, Waterhouse J. (2007). The influence of time awake and circadian rhythm upon performance on a frontal lobe task. *Neuropsychologia* 45:1966–1972.

Holyoak KJ. (1977). The form of analog size information in memory. *Cogn. Psychol.* 9:31–51.

Horne JA, Anderson NR., Wilkinson RT. (1983). Effects of sleep deprivation on signal detection measures of vigilance: implications for sleep function. *Sleep* 6:347–358.

Hut RA. (2007). CircWave. Version 1.4. Groningen, The Netherlands. Available at: http://www. euclock.org/modules.php?name=Content&pa=showpage&pid=20

Jasper I, Häußler A, Baur B, Marquardt C, Hermsdörfer J. (2009a). Circadian variations in the kinematics of hardwriting and grip strength. *Chronobiol. Int.* 26:576–594.

Jasper I, Häußler A, Marquardt C, Hermsdörfer J. (2009b). Circadian rhythm in handwriting. *J. Sleep Res.* 18:264–271.

Johnson MP, Duffy JF, Dijk DJ, Ronda JM, Dyal CM, Czeisler CA. (1992). Short-term memory, alertness and performance: a reappraisal of their relationship to body temperature. *J. Sleep Res.* 1:24–29.

Killgore WD, Balkin TJ, Wesensten NJ. (2006). Impaired decision making following 49 h of sleep deprivation. *J. Sleep Res.* 15:7–13.

Kristofferson MW. (1972). Effects of practice on character classification performance. *Can. J. Psychol.* 26:54–60.

Langner R, Steinborn MB, Chatterjee A, Sturm W, Willmes K. (2010). Mental fatigue and temporal preparation in simple-reaction time performance. *Acta Psychol.* 133:64–72.

Leth-Steensen C, Marley AAJ. (2000). A model of response time effects in symbolic comparison. *Psychol. Rev.* 107:62–100.

Lim J, Dinges DF. (2008). Sleep deprivation and vigilant attention. *Ann. N. Y. Acad. Sci.* 1129:305–322.

Link S. (1990). Modeling imageless thought: the relative judgment theory of numerical comparisons. *J. Math. Psychol.* 34:2–41.

Miccoli L, Versace F, Koterle S, Cavallero C. (2008). Comparing sleep-loss and sleep inertia: lapses make the difference. *Chronobiol. Int.* 25:725–744.

Monk TH, Carrier J. (1997). Speed of mental processing in the middle of the night. *Sleep* 20:399–401.

Moyer RS, Landauer TK. (1967). Time required for judgements of numerical inequality. *Nature* 215:1519–1520.

Nilsson JP, Soderstrom M, Karlsson AU, Lekander M, Åkerstedt T, Lindroth NE, Axelsson J. (2005). Less effective executive functioning after one night's sleep deprivation. *J. Sleep Res.* 14:1–6.

Oriet C, Tombu M, Jolicoeur P. (2005). Symbolic distance affects two processing loci in the number comparison task. *Mem. Cognit.* 33:913–926.

Poltrock SE. (1989). A random-walk model of digit comparison. *J. Math. Psychol.* 33:131–162.

Portaluppi F, Touitou Y, Smolensky M. (2008). Ethical and methodological standards for laboratory and medical biological rhythm research. *Chronobiol. Int.* 25:999–1016.

Roenneberg T, Wirz-Justice A, Merrow M. (2003). Life between clocks: daily temporal patterns of human chronotypes. *J. Biol. Rhythms* 18:80–90.

Rogers NL, Dorrian J, Dinges DF. (2003). Sleep, waking and neurobehavioural performance. *Front. Biosci.* 8:1056–1067.

Sagaspe P, Sanchez-Ortuno M, Charles A, Taillard J, Valtat C, Bioulac B, Philip P. (2006). Effects of sleep deprivation on Color-Word, Emotional, and Specific Stroop interference and on self-reported anxiety. *Brain Cogn.* 60:76–87.

Sanders AF. (1998). *Elements of human performance*. Mahwah, NJ: Lawrence Erlbaum.

Schmidt C, Collette F, Cajochen C, Peigneux P. (2007). A time to think: circadian rhythms in human cognition. *Cogn. Neuropsychol.* 24:755–789.

Sigman M, Dehaene S. (2005). Parsing a cognitive task: a characterization of the mind's bottleneck. *PLoS Biol.* 3:334–349.

Sternberg S. (1998). Discovering mental processing stages: the method of additive factors. In Scarborough D, Sternberg S. (eds.). *An invitation of to cognitive science: methods, models, and conceptual issues*. Vol. 4. Cambridge, MA: MIT Press, pp. 703–863.

Weber JM, Schwander JC, Unger I, Meier D. (1997). A direct ultrasensitive RIA for the determination of melatonin in human saliva: comparison with serum levels. *Sleep Res.* 26:757.

Wilkinson RT. (1990). Response-stimulus interval in choice serial reaction time: interaction with sleep deprivation, choice, and practice. *Q. J. Exp. Psychol.* 42A:401–423.