A Standard Procedure Enhances the Correlation Between Subjective and Objective Measures of Sleepiness

Chien-Ming Yang, PhD[1]; Fung-Wan Lin, BS1; Arthur J. Spielman, PhD[2]

[1]Fu Jen University, Department of Psychology, Taipei, Taiwan; [2]Sleep Disorders Center, New York Methodist Hospital, Brooklyn, NY

Objective: The goal of this study was to assess whether instituting a standard procedure to minimize transient activation prior to the subjective rating of sleepiness can improve the predictive value of the rating process.

Methods: Thirty young adults, aged 19 to 26 years, participated in the study. Subsequent to sleeping at home with bedtime restricted to 5 hours, they came to the sleep laboratory. They were instructed to rate their level of sleepiness on the Stanford Sleepiness Scale (SSS) and visual analog scales (VAS). A "calm-down" procedure, sitting quietly with eyes closed for 1 minute, was instituted prior to sleepiness ratings for half of the subjects (experimental group) but not for the other half of the subjects (control group). A nap trial with polysomnographic recording was then conducted, followed by a vigilance test.

Results: For the experimental group, VAS results of "sleepiness" and "alertness" both correlated significantly with sleep-onset latency during the nap (SOL: $r = -.62$ and $.64$, respectively, P values $< .05$) and with reaction time (RT) on the vigilance test ($r = .56$ and $-.54$, P values $< .05$). The SSS ratings showed significant correlation with nap SOL ($r = -.58$, $P < .05$) but not with RT on the vigilance test ($r = .19$, $p = .52$). For the control group, none of the subjective ratings showed significant correlation with objective measures. The differences between the resultant correlations for the 2 groups were statistically significant for 2 sets of correlations: the correlation between VAS of "alertness" and nap SOL and the correlation between VAS of "sleepiness" and RT on the vigilance test.

Conclusion: The results indicate that the subjective ratings of the sleepiness state for individuals with mild sleep restriction more faithfully reflect a physiologic tendency to fall asleep as well as cognitive attentiveness when the ratings are conducted subsequent to sitting still with eyes closed for a sufficient time to minimize transient activation.

Key Words: Sleepiness, subjective sleepiness, Stanford Sleepiness Scale, sleep-onset latency

Citation: Yang CM; Lin FW; Spielman AJ. A standard procedure enhances the correlation between subjective and objective measures of sleepiness. *SLEEP* 2004;27(2):329-32.

INTRODUCTION

SLEEPINESS, DEFINED AS BOTH A FUNDAMENTAL FEELING STATE AND THE TENDENCY TO FALL ASLEEP,[1] IS ONE OF THE MOST IMPORTANT CONCEPTS IN SLEEP MEDICINE. At a given moment, the manifest feeling of sleepiness may reflect basal physiologic sleepiness as well as factors such as context, activity, and motivation. When these activating influences are absent for an individual, latent or physiologic sleepiness is revealed. It has been suggested that subjective ratings of sleepiness reflect manifest sleepiness and objective measures such as sleep-onset latency (SOL) and behavioral measures reflect latent sleepiness.[1] The significant but limited correlations between manifest and latent measures of sleepiness suggest that these different assessment techniques do not always measure similar aspects of sleepiness. The correlation between subjective sleepiness ratings and sleep latencies on the Multiple Sleep Latency Test (MSLT), for example, was initially found to be high in normal subjects[2,3] but was subsequently found to be nonsignificant in both normal subjects[4-7] and clinical populations.[7-9]

Similarly, the correlations between subjective ratings and objective behavioral measures of sleepiness have also been shown to have limited correspondence. For example, one study reported small to moderate correlations (r values = .25 to .42) between subjective sleepiness rating on visual analog scales (VAS) and performance on auditory vigilance tasks following sleep reduction, when data points collected repeatedly in 8 subjects were all included.[10] Another study showed moderate to high correlations (r values = .49 to .71) between performance on a visual vigilance task and subjective sleepiness ratings on VAS and a 9-point rating scale (Karolinska Sleepiness Scale), when the 48 data points collected repeatedly from 6 subjects were included.[11] Furthermore, the Stanford Sleepiness Scale (SSS) has been reported to have moderate to high correlations (r values = .47 to .70) with performance on different cognitive tasks. The correlations ranged from .18 to .95 for different subjects.[12] A recent study demonstrated that mild physical activity could mask subjective sleepiness following total sleep deprivation. Subjects showed less decline of subjective sleepiness after walking, but their cognitive performance showed the same degree of impairment regardless of their having participated in exercise prior to subjective-sleepiness rating.[13]

As activating situational factors are minimized, manifest sleepiness should better correspond to measures of latent sleepiness. The current study was designed to minimize the situational factors that counter latent sleepiness and weaken the association between subjective and objective measures of sleepiness.

The present study was conducted to clarify the relationship between objective and subjective sleepiness measures. The goals of the study were (1) to evaluate the relationship between subjective sleepiness and physiologic tendency to fall asleep measured by polysomnographic SOL, (2) to evaluate the relationship between subjective sleepiness and vigilance measured by a continuous performance task, and (3) to determine if minimizing transient activation by instituting a procedure to calm the subject down prior to subjective-sleepiness ratings improves the predictive value of the rating process.

METHOD

Subjects

Subjects were 30 young adults (10 men and 20 women) recruited from a college campus. Their ages ranged from 19 to 26 years, with a mean age of 22 years and SD of 1.36 years. Their habitual bedtime ranged from 11:00 PM to 2:00 AM, and habitual arising time from 7:30 AM to 11:30 AM. Their habitual total time in bed ranged from about 6 hours to

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

Submitted for publication March 2003
Accepted for publication October 2003
Address correspondence to: Chien-Ming Yang, PhD, Fu Jen University, Department of Psychology, 510 Chung Cheng RD, Hsin Chuang, Taipei Hsien, Taiwan 24205; Tel: 886-2-29031111 ext. 2131; Fax: 886-2-29010171; E-mail: yang_cm@yahoo.com

9 hours. After passing a screening interview for psychiatric disorders, neurologic disorders, and major medical conditions, the subjects were randomly assigned to either the experimental group or the control group.

PROCEDURE

The day before coming to the laboratory, the subjects were instructed to refrain from consuming caffeinated drinks. Their sleep was restricted to 5 hours (from 3:00 AM to 8:00 AM) to induce mild sleepiness. They were instructed to call a time-stamping answering machine in the laboratory immediately prior to their going to bed and immediately subsequent to their having woken up. The subjects were scheduled to come to the sleep laboratory at 9:00 AM, 2:00 PM, or 6:00 PM, at the subjects' convenience. Eight subjects were scheduled for the morning session (3 in control group, 5 in experimental group); 8 for the afternoon session (6 in control group, 2 in experimental group); 14 for the evening session (6 in control group, 8 in experimental group). The habitual bedtime and rising time did not differ significantly among the subjects in the morning, afternoon, and evening groups (bedtime: F = .01, NS; arising time: F = 2.0, NS).

The laboratory was sound attenuated and air-conditioned. Indoor fluorescent lights were on throughout the experimental procedures except for the nap test. During the nap test, all lights were turned off.

After the subjects arrived in the laboratory, electrodes were applied for standard polysomnographic recording. The recording montage included electroencephalography (C3, C4, O1, and O2 referred to A1 and A2), electrooculography, and chin electromyography. Two sessions of sleepiness measures were conducted. The first was an adaptation session. Only data from the second session were included for analysis. During each session, the subjects' subjective sleepiness was measured with the SSS and VAS. The SSS is a 7-point scale ranging from 1 ("Feeling active and vital; alert; wide awake.") to 7 ("Almost in reverie; sleep onset soon; lost struggle to remain awake.").[12] The VAS contained 2 items ("sleepiness" and "alertness"), each consisting of a 10-cm line, with "not sleepy/not alert at all" on the left end of the line and "extremely sleepy/extremely alert" on the right end of the lines. Subjects were asked to view the line as representing their personal range of feelings and to place a mark on the line indicating their feeling at that moment. Similar scales have been used in many of the previously mentioned studies.[10,11,13] Subjects in the experimental group were instructed to sit quietly with their eyes closed for 1 minute before the ratings. No specific instruction was given to subjects in the control group. After the ratings, subjects were instructed to lie down in bed and try to fall asleep. As soon as their polysomnogram showed 3 consecutive pages of sleep following the standard scoring criteria,[14] they were awakened and the recording terminated. If sleep onset was not achieved, the recording was terminated in 25 minutes. A continuous performance test (CPT) measuring vigilance was then administered. The CPT was modified from the Multiple Vigilance Test[15] and consisted of 160 stimuli presented on a computer screen. The interstimulus intervals were varied randomly between 4 and 11 seconds. Forty of the stimuli were targets, and 120 were nontargets. Subjects were instructed to hit the space bar as soon as possible when the target stimulus was presented and to not respond to nontargets. There was a 20-minute break between the adaptation session and formal session. During the break, the subjects were allowed to go to the restroom. For the rest of the time, the experimenters engaged in casual conversation with the subjects.

RESULTS

Polysomnography SOL

All subjects fell asleep within 25 minutes. The SOL, defined by the time from lights out to the first of 3 continuous epochs of sleep, was used as a measure of objective sleepiness. The mean SOLs were 10.5 (± 5.3) minutes for control subjects and 8.2 (± 3.5) minutes for experimental subjects (t = 1.40, NS). These values were within the expected range following mild sleep deprivation. The mean SOLs obtained from the 3 different times of day (morning = 8.6 ± 4.1; afternoon = 9.9 ± 5.2; evening = 9.4 ± 4.7) were also compared with 1-way analysis of variance and were not statistically different (t = 0.16, NS).

Pearson correlations were conducted between SOL and the sleepiness ratings. Both sleepiness ratings correlated significantly with SOL for the experimental group. However, for the control group, neither of the sleepiness ratings correlated with SOL (see Table 1). Figure 1 shows the scatter plots of the correlations. The differences between the respective correlations for the 2 groups were also compared by computing the confidence interval corresponding to the z-score differences. The results revealed that only the correlation between VAS alertness rating and SOL was significantly higher for the experimental group than the control group ($P < .05$). The remainder of the correlations revealed no statistical differences between the 2 groups.

Continuous Performance Test

Reaction time (RT) for correct response to targets, hit rate (number of correct responses to a target divided by the total number of targets), and number of commission errors (responding to nontarget stimuli) and lapses (responses with an RT longer than 3 times the grand mean) were calculated for analysis. One of the subjects in the experimental group obtained a mean RT of 1.74 seconds, which is more than 10 SD above the mean RT of the rest of the subjects (mean = .62, SD = .11 second). Her data on the CPT were therefore excluded from data analysis. For the experimental group, RTs correlated significantly with VAS sleepiness and alertness ratings but not with SSS ratings (see Table 1). Figure 2 shows the scatter plots of the correlations. Hit rate, commission errors, and lapses on the CPT did not correlate with any of the sleepiness ratings. Again, for the control group, none of the CPT measures correlated significantly with any of the sleepiness ratings. Comparisons of the resultant correlations between experimental and control groups were also conducted by computing the confidence interval on the z-score differences, results showing that only the correlation between VAS sleepiness rating and RT was significantly higher for the experimental group than the control group ($P < .05$). The remainder of the correlations did not differ statistically between the 2 groups.

DISCUSSION

The present study revealed that a simple "calm-down" procedure conducted prior to subjective ratings of sleepiness could enhance the correlation between an individual's subjective sleepiness and that individual's SOL and performance on a vigilance task. The standard procedure increased the predictive value of the subjective ratings of sleepiness from a nonsignificant correlation to a moderately high correlation with objective measures of sleepiness. When comparing the correlations obtained with and without the standard procedure, two of the comparisons reached significance as regards difference.

Table 1—Correlations between subjective ratings of sleepiness states and sleep-onset latency and performance on continuous performance test

		Continuous Performance Test			
	SOL	RT	Hit	CE	Lapse
Experimental Group					
SSS	-.578*	.188	-.130	.123	.225
VAS-Sleepiness	-.616*	.561*	-.114	.171	.179
VAS-Alertness	.640*	-.541*	.077	-.071	-.194
Control Group					
SSS	-.353	-.293	-.397	.313	-.383
VAS-Sleepiness	-.191	-.038	-.462	-.046	-.192
VAS-Alertness	.049	-.105	.273	-.093	-.430

*$P < .05$

SOL refers to sleep-onset latency; RT, reaction time; Hit, hit rate; CE, commission errors; Lapse, number of lapses; SSS, Stanford Sleepiness Scale; VAS, visual-analog scale.

One possible explanation for our finding that the experimental group's ratings of subjective sleepiness corresponded more satisfactorily to SOL and cognitive attentiveness is that sitting quietly for 1 minute with eyes closed reduces situational factors that produce activation and therefore unmasks sleepiness and sleep tendency. By reducing activation due to motivation and sensory and motor stimulation, the experience of sleepiness is more in accord with behavioral and physiologic indexes of sleepiness.

From a different vantage point, the current results may be understood as being due to reducing the method variance between our sleepiness measures. Testing theory has long held that different procedures used to measure a construct will produce different results in part due to differ-

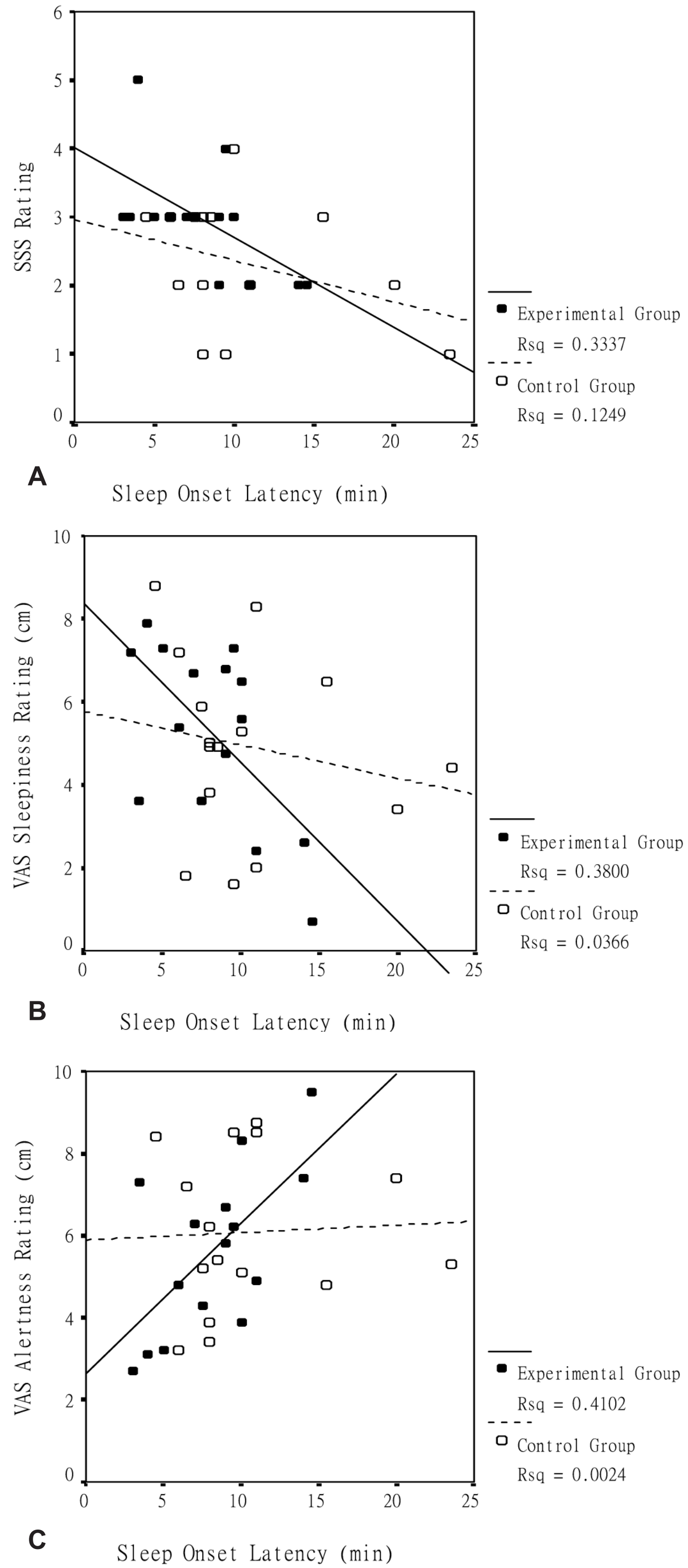

Figure 1—Scatter plots of the correlations between subjective ratings of sleepiness and sleep-onset latency (SOL) during the nap opportunity. (a) Stanford Sleepiness Scale (SSS) rating versus SOL during nap the opportunity; (b) visual-analog scale (VAS) sleepiness rating versus SOL during nap the opportunity; (c) VAS alertness rating versus SOL during nap the opportunity.

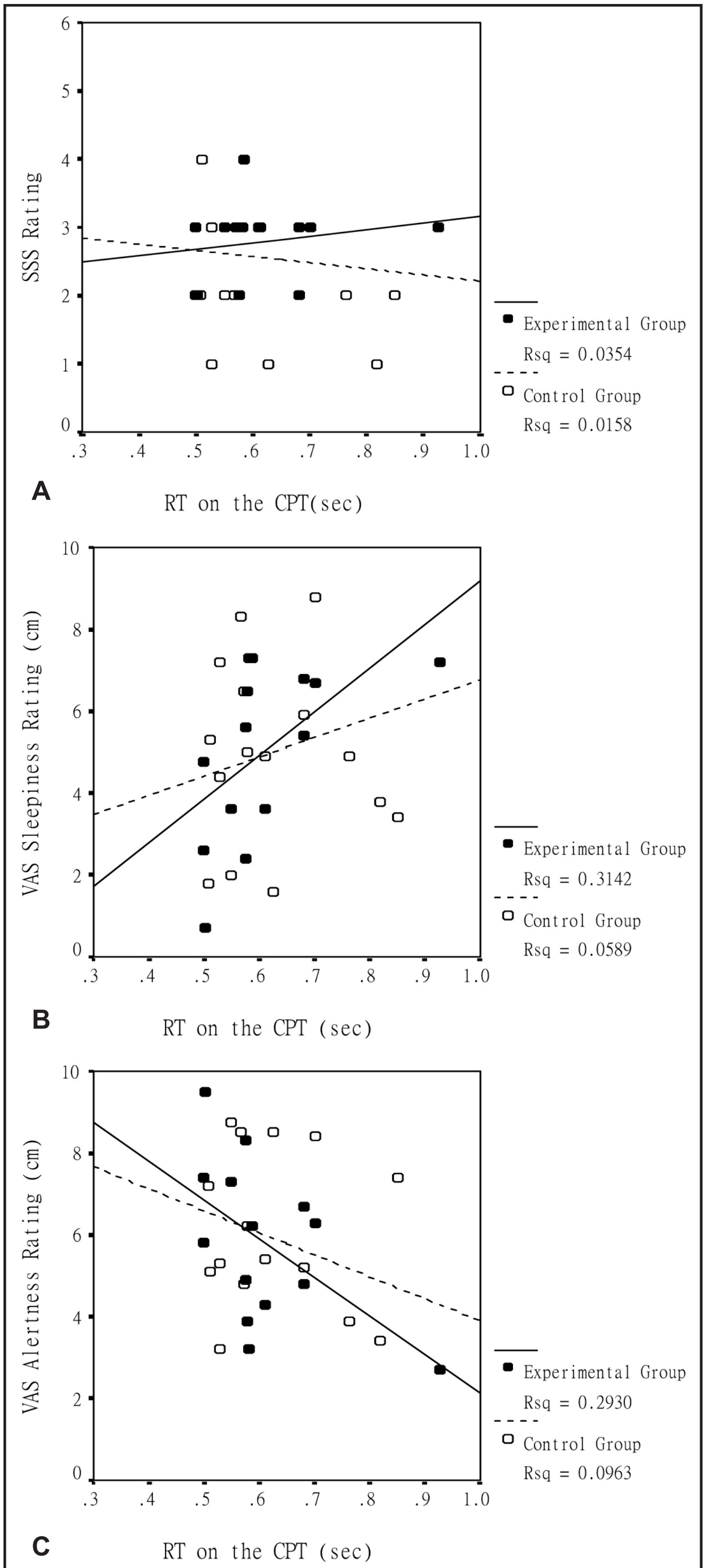

Figure 2—Scatter plots of the correlations between subjective ratings of sleepiness and reaction time (RT) on the continuous performance test (CPT). (a) Stanford Sleepiness Scale (SSS) rating versus RT on the CPT; (b) visual-analog scale (VAS) sleepiness rating versus RT on the CPT; (c) VAS alertness rating versus RT on the CPT.

ences in the methods themselves.[16] This principle was supported by the results of the present study, which demonstrated that when the procedure prior to ratings of subjective sleepiness was standardized so that it resembled the procedure prior to a nap, subjective sleepiness more closely reflected the physiologic tendency to fall asleep. The standard instructions to the experimental group—sit quietly for 1 minute with eyes closed—is similar to the procedure prior to the nap. We speculate that if the procedures prior to rating subjective sleepiness were even more similar to the prenap procedure, for example, lying down in the dark with eyes closed, the correspondence between these two measures would be even stronger. We decided that subjects would sit up rather than lie down so that this standard procedure could be widely applicable in other settings when a bedroom is not available. If the investigator or clinician is interested in the subjective experience of sleepiness that is less related to momentary activating factors and more related to behavioral and physiologic sleepiness, we suggest that the current procedures become a standard.

This study was conducted in individuals with mild restriction of sleep the night before the testing. Future studies will be needed to determine if the standard procedures for subjective rating of sleepiness we recommend improve predictions of objective measures of sleepiness in more sleep-deprived (eg, total sleep-deprived), non–sleep-deprived, and sleep-disordered individuals. In addition, the procedure may be modified for different conditions. For example, eyes being held closed for a period of 1 minute may be too long for totally sleep-deprived individuals because they may actually fall asleep during this procedure. On the other hand, 1 minute may not be sufficiently long subsequent to active physical activity. Also, it has been shown that subjective rating of sleepiness following the completion of a performance task yields better predictive value of cognitive performance than was the case if the rating process was completed prior to the performance task.[17] The performance task itself might serve as a standard procedure to minimize the transient activation. Therefore, the necessity for the standard procedure prior to subjective rating of sleepiness may also depend upon the activity conducted immediately prior to the rating process.

REFERENCES

1. Dement WC, Carskadon MA. Current perspectives on daytime sleepiness: the issues. Sleep 1982;5:S56-66.
2. Carskadon MA, Dement WC. Effects of total sleep loss on sleep tendency. Percept Mot Skills 1979;48:495-506.
3. Carskadon MA, Dement WC. Cumulative effects of sleep restriction on daytime sleepiness. Psychophysiology 1981;18:107-13.
4. Johnson LC, Freeman CR, Spinweber CL, Gomez SA. The relationship between subjective and objective measures of sleepiness. Psychophysiology 1991;28:65-71.
5. Sugerman JL, Walsh JK. Physiological sleep tendency and ability to maintain alertness at night. Sleep 1989;12:106-12.
6. Richardson GS, Carskadon MA, Orval EJ, Dement WC. Circadian variations of sleep tendency in elderly and young adult subjects. Sleep 1982;5:S82-94.
7. Seidel WF, Ball S, Cohen S, Patterson N., Yost D, Dement WC. Daytime alertness in relation to mood, performance, and nocturnal sleep in chronic insomnia and noncomplaining sleepers. Sleep 1984;7:230-8.
8. Roth T, Hartse KM, Zorick F, Conway W. Multiple naps and the evaluation of daytime sleepiness in patients with upper airway sleep apnea. Sleep 1980;3:425-39.
9. Sugerman JL, Stern JA, Walsh JK. Daytime alertness in subjective and objective insomnia: Some preliminary findings. Biol Psychiatry 1985;20:741-50.
10. Devoto A, Lucidi F, Violani C, Bertini M. Effects of different sleep reductions on daytime sleepiness. Sleep 1999;22:336-43.
11. Gillberg M, Kecklund G, Akerstedt T. Relations between performance and subjective ratings of sleepiness during a night awake. Sleep 1994;17:236-41.
12. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: A new approach. Psychophysiology 1973;10:431-6.
13. Matsumoto Y, Mishima K, Satoh K, Shimizu T, Hishikawa Y. Physical activity increases the dissociation between subjective sleepiness and objective performance levels during extended wakefulness in human. Neurosci Lett 2002;326:133-6.
14. Rechtschaffen A, Kales A. A Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Subjects. Washington, DC: Public Health Service, US Government Printing Office; 1968.
15. Hirshkowitz M, De La Cueva L, Herman JH. The multiple vigilance test. Behav Res Method, Instrum Comp 1993;25:272-5.
16. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 1959;56:81-105.
17. Van Dongen HPA, Dinges DF. Circadian rhythms in fatigue, alertness, and performance. In: Kryger MH, Roth T, Dement WC, eds. Principles and Practice of Sleep Medicine, 3rd ed. Philadelphia: WB Saunders Company; 2000:391-9.