Proceedings of the

37th Annual Conference of the International Military Testing Association

IMTA

TORONTO '95

Toronto, Ontario, Canada
16 - 19 October 1995

hosted by the
Canadian Forces Personnel Applied Research Unit

DCIEM NO. 95-41

Evidence for a Modafinil Induced "Overconfidence" Effect During Sustained Operations

Joseph V. Baranski
Defence and Civil Institute of Environmental Medicine, North York, Ontario, Canada

"Self-monitoring" refers to the ability to assess accurately one's own performance in a specific environment. The present paper explored the effects of the alerting substances Modafinil and Amphetamine on the ability to self-monitor cognitive performance during 64 hours of sleep deprivation (SD) and continuous cognitive work. Two tasks were examined for self-monitoring ability: A visual (perceptual) comparison task and a complex mental addition task. Immediately before [*and after*] each session, subjects were required to estimate the proportion of trials that they thought they would answer [*had answered*] correctly. The experiment involved three drug conditions: Modafinil (300 mg, N=14), Amphetamine (20 mg, N=12), and Placebo (N=13), which were administered three times: (1) at 2330 prior to the first night of SD, (2) at 0530 of the second night of SD, and (3) at 1530 of the third day of SD. The placebo condition displayed substantial performance deficits (i.e., poorer response accuracy) during the first and second nights of SD. Concomitantly, however, the self-monitoring ability of the placebo condition was remarkably accurate throughout the study. The effect of the stimulants was to effectively maintain task performance through the first night of SD and to recuperate performance during the second night of SD. However, whereas Amphetamine did not disturb the ability to self-monitor cognitive performance during SD, subjects in the Modafinil condition displayed a distinct, short-term "*over-confidence*" effect (i.e., an overestimation of actual cognitive capability), two-hours post-dose.

INTRODUCTION

Recent international conflicts such as 'The Gulf War' have provided striking examples of how major advances in contemporary military technology now permit 'round-the-clock' or sustained operations (SUSOPS). In sharp contrast, the human element embedded in this complex technological assembly remains severely challenged by evolution: We must sleep. The inevitable consequence of this discrepancy in evolutionary speed has been a renewed and burgeoning interest in the human factors of SUSOPS.

As an example, recent initiatives such as the U.S. Tactical Air Command "Aircrew Conditioning Program" (see Emonson & Vanderbeek, 1995), employed in both Desert Storm and Desert Shield, have highlighted the importance of a generalized approach to the effective management of human fatigue. A component of this program was the explicit recommendation to U.S. Air Force personnel to use d-amphetamine (5 mg) "30 minutes before critical stages of flight if they felt unduly fatigued (Emonson & Vanderbeek, 1995, p.260)." Although the effectiveness of d-amphetamine to combat the debilitating effects of fatigue is now well documented, it is also the case that amphetamine possesses several highly undesirable properties (e.g., cardiovascular side effects, insomnia, tolerance, and addiction). More recently, a stimulant called Modafinil (Lafon Labs, France) has been intensively studied as a potential alternative to d-amphetamine (see Lyons & French, 1991). The drug has been reported to be as effective as d-amphetamine but without many of the undesirable side effects (see Buguet et al., in press), and has in fact been used by the French military during The Gulf War.

In the Winter of 1994, under the auspices of the Franco-Canadian Research Accord, the Human Factors division of DCIEM undertook an extensive investigation of the effectiveness of Modafinil (as compared to d-amphetamine and placebo) in a simulated SUSOPS environment involving normal healthy adults. Generally, the principle empirical findings concerning Modafinil have been quite encouraging (see Buguet et al., in press; Pigeau et al., in press). In the present paper, we present additional results emanating from the above mentioned study. Here, we examine the effects of Modafinil, d-amphetamine, and placebo on the ability to self-monitor cognitive performance during a period of extensive sleep deprivation (SD) and sustained mental work.

The term *self-monitoring* is used here to denote the global, "meta-cognitive" ability to assess accurately one's own performance in a given environment. To date, investigations of the ability to self-monitor performance have typically involved an assessment of the validity of subjective *confidence* ratings made for individual judgments within a specific cognitive task (for reviews, see Baranski & Petrusic, 1994, in press; Baranski et al., 1994; Keren, 1991; Lichtenstein et al., 1982). Alternatively, one can assess the validity of subjective estimates of overall task performance (Baranski & Pigeau, 1995; Gigerenzer et al., 1991; Stone, 1994). Central to this issue is that many fundamental cognitive abilities are known to display circadian rhythms and progressive deterioration with increasing SD (Babkoff et al., 1991; Hockey, 1986; Kreuger, 1989). Consequently, it is important to understand the extent to which the sleep deprived individual can quantify and articulate these cognitive performance deficits and the extent to which stimulating drugs such as Modafinil and d-amphetamine affect the relationship between subjective and overt performance.

METHOD

Subjects. Forty-one Canadian Forces personnel (39 M) participated for six consecutive days in exchange for basic pay plus stress allowance.

Procedure. A detailed description of the general methodology is provided in Pigeau et al. (1995). Briefly, subjects ingested the same drug (Modafinil, 300 mg; d-Amphetamine, 20 mg; or placebo) three times during the course of the study: (1) 2330 h prior to the first night of sleep deprivation; this was to determine if the stimulating drugs could prevent the well documented "first night" dip in performance, (2) 0530 h of the second night of SD; this was to determine if the stimulating drugs could recuperate performance, and (3) 1530 h of the third day of SD; this was primarily to examine the effect of the stimulating drugs on recovery sleep onset (see Buguet et al., 1995).

Upon arrival (Sunday), subjects were briefed on the experimental protocol and were given short practice sessions on the battery of cognitive tasks to be used in the experiment. More formal training (i.e., baseline) sessions resumed on Monday (1200 h - 1800 h). Subjects were awakened at 0600h on Tuesday and immediately began the SUSOPS portion of the experiment, which continued until 1800h on Thursday. On Thursday night, subjects retired for a maximum of 13 hours of recovery sleep. The subjects then participated for a final recovery session on Friday (1200 h -1800 h). Following a second night of recovery sleep, the subjects were debriefed and released at approximately 1000h on Saturday.

Throughout the SD period the subjects performed continuous cognitive work in 2-hour sessions; 1-hour and 45 minutes of work followed by a 15 minute break. EEG, ECG, blood pressure, and core and surface body temperatures were recorded throughout the study. Subjects performed a variety of cognitive tasks (see Pigeau et al., 1995) but we focus here on the two tasks that were used to assess the effects of the stimulating drugs on the ability to self-monitor cognitive performance during SUSOPS conditions.

Perceptual Comparison Task. The perceptual comparison task required relative judgments of line length. The subjects task was to determine the longer or shorter of two horizontal lines presented on a PC monitor. Subjects responded by depressing the left or the right button on a PC mouse. Line length pairs were chosen so as to provide an intermediate level of judgment difficulty (Baranski & Petrusic, 1992). Subjects completed two 4.5 minute sessions of the comparison task in each two-hour block in the study. Prior to beginning each session, the subjects were prompted for a *pre-task estimate* of performance. Their task was to type in the proportion of trials that they thought *they would answer* correctly. Immediately following the completion of each session, subjects were prompted for a *post-task estimate* of performance. Here their task was to type in the proportion of trials they thought *they answered* correctly. It was emphasized that 100% was to denote that all trials were (would be) answered correctly and because there were only two choices on each trial, 50% correct could be achieved by chance responding.

Mental Addition task. The addition task required subjects to add a sequence of 8 randomly presented numbers (1-16), which were presented on a computer monitor at a rate of one number every 1.25 seconds. The sequence was terminated by the presentation of a visual prompt (=>) at which time subjects typed in their response and then pressed the 'Enter' key. Subjects performed one 15-minute session of the addition task in each of the 12 six-hour blocks in the study. As with the comparison task, subjects provided pre- and post-task estimates of performance for the mental addition task. Here, it was emphasized that 100% was to denote that all trials were (will be) answered correctly and that 0% was to denote that all trials were (will be) answered incorrectly. In other words, it was clear to the subjects that a "half-range" probability scale should be used for the comparison task (i.e., 50%-100%) and a "full-range" probability scale should be used for the addition task (i.e., 0%-100%).

RESULTS

The data of two subjects were excluded from the analyses of the perceptual task because their performance remained at approximately 50% correct throughout the study (i.e., perhaps because of a visual impairment, these subjects could not do the task).

Perceptual Comparison Task

Response accuracy. Figure 1 provides a plot of the accuracy of perceptual comparisons as a function of SD, separately for each of the three drug conditions. We provide an index of performance in terms of the difference in the percentage of correct responses from baseline (Monday).

The data for the placebo group displays the expected effects of extensive SD on performance; a steady decline during the first night, a leveling-off of performance during the second day, a second decline during the second night, and a full recovery following a restorative sleep. For the Amphetamine group, the first drug administration effectively maintained performance through the first night, there was a rapid decline in performance during the second night because the effect of the drug had worn off, and there was a restorative effect due to the second drug administration. For the Modafinil group, there was also a rapid decline in performance during the second night and a restorative effect of the second drug administration; generally, performance in the Modafinil condition was superior to that for the placebo group but not as good as that for the Amphetamine group. These results were confirmed by ANOVA.

The main effect of SESSIONS was highly reliable ($F(30, 1020)=5.34$, $p < .0001$). Trend analyses [linear ($p<.0009$), quadratic ($p<.005$), and cubic ($p<.003$)] confirmed the general decline in

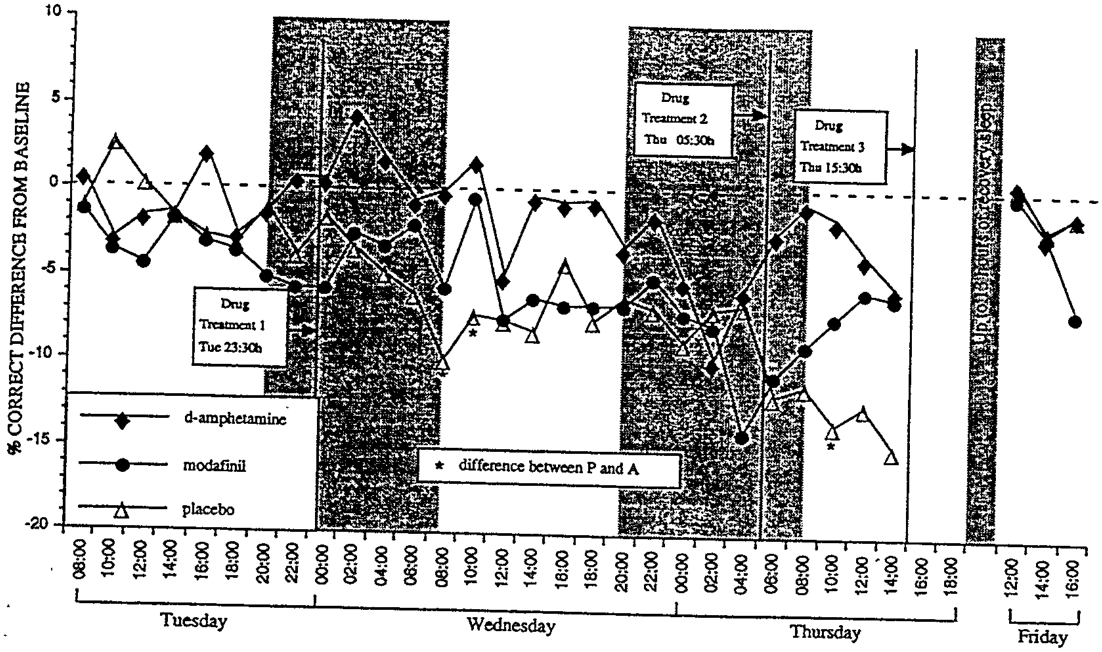

Figure 1. Percent correct difference from baseline for the three drug conditions in the perceptual task.

performance and the circadian effects, respectively. Most importantly, the interaction between DRUG condition and SESSIONS was reliable ($F(60,1020) = 1.74$, $p < .025$), confirming that performance during the SD period was differentially affected by the drug manipulation. Newman-Keuls post-hoc analyses revealed that performance in the Amphetamine group was superior to that of the placebo group for each of the 4 sessions indicated by an asterisk in Figure 1 (i.e., 0800 and 1000 on Wednesday and Thursday).

Long- and short-term drug effects on overt and subjective performance. The difference between the pre- and post-task estimates of performance and the actual proportion of correct responses provided the index of self-monitoring ability. Briefly, the results revealed that all groups displayed generally accurate assessments of performance throughout the SD period. On the other hand, a specific aspect of the data encouraged more detailed analyses. Most notably, the Modafinil condition displayed a distinct overconfidence immediately following each drug administration. This pattern was not evident in the placebo or Amphetamine conditions.

An increase in overconfidence can occur in two ways: performance declines but estimates do not (i.e., the person does not realize that they are doing

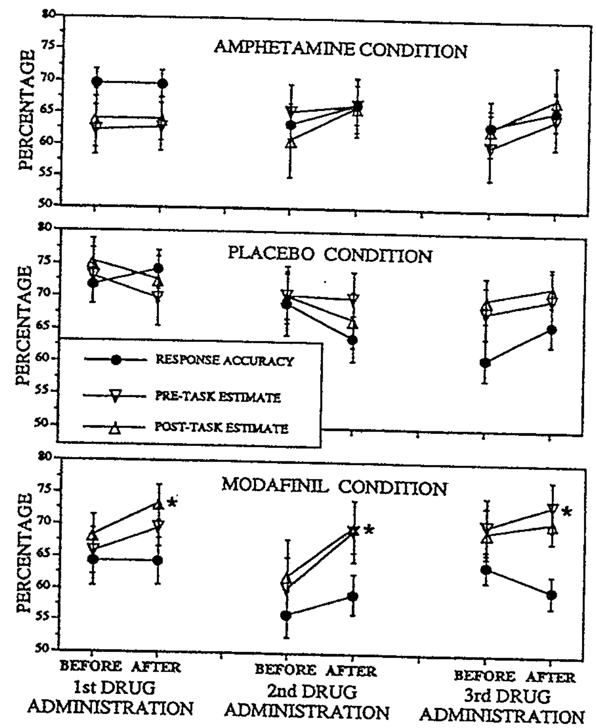

Figure 2. Percentage of correct responses and Pre- and Post-task estimates immediately before and after the three drug administrations in the perceptual comparison task.

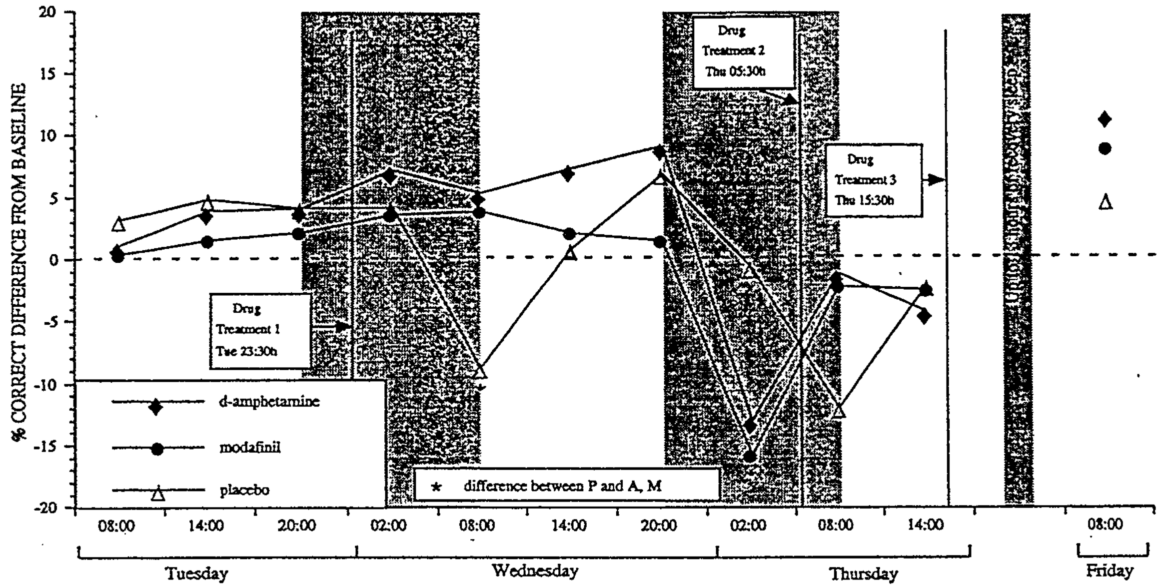

Figure 3. Percent correct difference from baseline for the three drug conditions in the mental addition task.

worse) or estimates increase but performance does not (i.e., the person thinks they are doing better than they actually are). Figure 2 provides plots of the proportion correct and estimated proportion correct immediately before and immediately after each drug administration.

The data for the Amphetamine and placebo conditions display generally accurate assessments of performance; all estimates were not significantly different from actual performance. In the Modafinil condition we see a different picture. Following the first administration, performance does not change but post-task estimates increase (p=0.05) and the increase in the pre-task estimate approached significance. Following the second administration, performance showed a non-significant increase but the pre-task and the post-task estimates increased, resulting in reliable overconfidence in the pre-task estimate ($p < .044$) and the post-task estimate ($p < .032$). Finally, following the third administration, performance actually declined by 5% but subjective estimates rose by over 10%, resulting in a substantial and highly significant degree of overconfidence.

Summary. Performance on the perceptual comparison task was clearly effected by SD; generally, accuracy decreased as SD increased. In terms of the effects of the drugs on performance, Amphetamine was superior to Modafinil which was superior to placebo. In terms of the ability to self-monitor cognitive performance, the global effect of SD was minimal and there were no consistent effects of the drug conditions. However, subjects under the influence of Modafinil displayed a clear overconfidence effect immediately following the ingestion of Modafinil (i.e., 2 hours post-dose). The generality of these findings is examined below.

Mental Addition Task

Response accuracy. Figure 3 provides plots of the percentage of correct additions for each of the drug conditions. For clarity, the data are once again presented in terms of a difference from baseline performance (Monday). As expected, the placebo condition shows performance dips during the first and second nights. The data for the Modafinil and the Amphetamine conditions paralleled one another quite closely: Performance in these conditions was maintained through the first night, there was a large dip in performance when the drugs wore off during the second night, and there was a clear restoration of performance following the second drug administration. All of these results were confirmed by ANOVA.

The main effect of SESSIONS was reliable ($F(10,360)=10.20$, $p < .0001$). Trend analyses revealed reliable linear ($p<.05$), quadratic ($p<.05$), and cubic ($p<.0001$) components. In addition, the interaction between SESSION and DRUG condition was also highly reliable, $F(20, 360)=2.28$, $p < .008$). Newman-Keuls post hoc analyses confirmed that the Modafinil and Amphetamine groups outperformed the placebo group at 0800 on Wednesday. However, the placebo group outperformed the Modafinil and the Amphetamine groups at 0200 on Thursday. The latter occurred because of a more rapid decline in performance for the Modafinil and Amphetamine conditions during the second night, after the effects of the drugs had worn off.

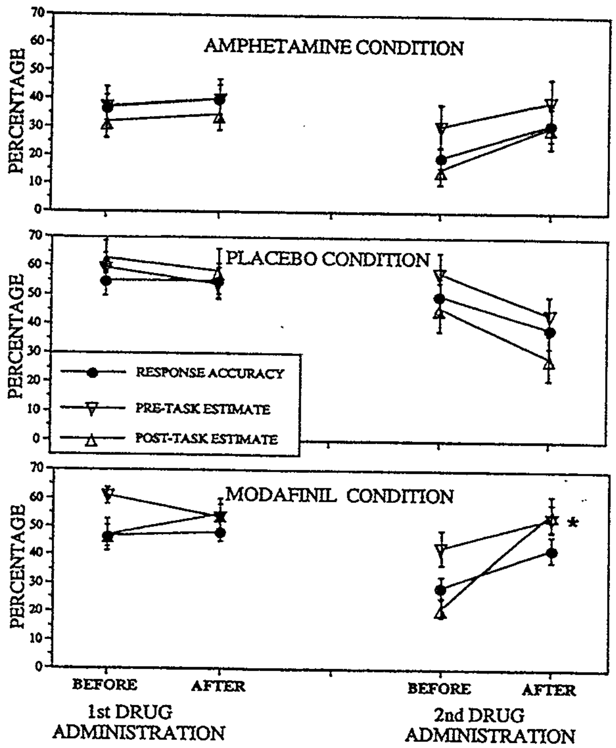

Figure 4. Percentage of correct responses and Pre- and Post-task estimates immediately before and after the two drug administrations in the mental addition task.

Long- and short-term drug effects on overt and subjective performance. As in the perceptual comparison task, all groups displayed generally accurate assessments of performance throughout the SD period. In addition, only the Modafinil condition exhibited increased overconfidence immediately following a drug administration. In Figure 4 we provide plots comparable to those reported in Figure 2 for the comparison task (note that the addition task was not scheduled following the third drug administration). As before, the Amphetamine and placebo conditions show accurate pre- and post-task estimates immediately following both drug administrations. For the Modafinil condition, however, there was a 7.5% increase in the post-task estimate following the first drug administration with no change in overt performance (this difference approached statistical significance). Following the second drug administration, a 13.7% increase in performance was accompanied by a highly significant 33.6% increase in the post-task estimate. Thus, here again, the immediate effect of Modafinil was to induce an exaggerated subjective estimate of performance.

Summary. Performance on the mental addition task was effected by SD. The placebo condition displayed the expected decline in performance during the first and second nights. Modafinil and Amphetamine worked equally well to maintain performance during the first night and to recuperate performance following the second night. On the other hand, the ability to self-monitor performance was quite accurate and showed minimal circadian effects. However, as in the comparison task, Modafinil appears to have an immediate, short-term influence on the ability to self-monitor performance. This intriguing and potentially important result clearly demands further study.

DISCUSSION

In a recent paper, Warot et al. (1993) compared the subjective profile of Modafinil with that of d-amphetamine, caffeine, and placebo (the dose of Modafinil was the same as that used in the present study). The authors reported that "Modafinil showed no sedation (PCAG scale), no pronounced elation or euphoria (MBG scale), *increased sensation of energy and intellectual efficiency* (BG scale) and very few somatic or dysphoric effects (LSD scale)" (Warot et al., 1993, p. 206, italics added). Our interest in the Warot et al. findings concern the subjective increase in intellectual efficiency while under the influence of Modafinil since it provides a possible link to the present findings of an immediate, short-term overconfidence effect. Indeed, a review of their Figure 1 reveals a Modafinil induced increase on the BG (Benzedrine Group) scale, which reaches a maximum at 2-hours post dose; at 4-hours post-dose the effect had returned to baseline. Hence, this may explain why our finding of overconfidence for Modafinil was immediate but short lived; the subjective increase in intellectual efficiency had probably disappeared by the time of our next experimental block.

In conclusion, the present results, together with those reported by Pigeau et al. (1995) and Buguet et al. (1995), highlight the potential utility of stimulating drugs in sustained operational settings. There is no doubt that cognitive performance can be maintained at a higher level, and for a longer duration, with stimulating drugs than without. In addition, Modafinil should be considered as a viable alternative to Amphetamine as it appears to posses some desirable properties (e.g., more rapid sleep onset, lower susceptibility to addiction, and fewer physiological side effects). However, as with any stimulant, Modafinil induces subjective effects. When such effects are consistent with overt performance, then there is much to benefit in terms of increased performance and productivity. On the other hand, if subjective effects are not consistent with overt performance, then the potential for human error necessarily increases.

This study was supported by a grant from Laboratiore L. Lafon. I would like to thank Andrea Hawton and Chad Berry for invaluable technical assistance. Correspondence: Dr. Joseph V. Baranski, Defence and Civil Institute of Environmental Medicine, 1133 Sheppard Ave. West, North York, Ontario, CANADA, M3M 3B9.

REFERENCES

Babkoff, H., Caspy, T., Mikulincer, M., & Sing, H. (1991). Monotonic and rhythmic influences: A challenge for sleep deprivation research. *Psychological Bulletin*, **109**, 411-428.

Baranski, J. V. & Petrusic, W.M. (1992). The discriminability of remembered magnitudes. *Memory & Cognition*, **20**, 254-270.

Baranski, J. V. & Petrusic, W.M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics*, **55**, 412-428.

Baranski, J. V. & Petrusic, W.M. (In Press). On the calibration of knowledge and perception. *Canadian Journal of Experimental Psychology*.

Baranski, J. V. & Pigeau, R. (1995). Self-Monitoring cognitive performance during sleep deprivation: Effects of modafinil, amphetamine, and placebo. *Sleep Research*, **24A**, 421.

Baranski, J. V., Pigeau, R. & Angus, R. (1994). On the ability to self-monitor cognitive performance during sleep deprivation: A calibration study. *Journal of Sleep Research*, **3**, 36-44.

Buguet, A., Montmayeur, A., Pigeau, R. & Naitoh, P. (In Press). Modafinil, d-Amphetamine and placebo during 64 hours of sustained mental work: II. Effect on two nights of recovery sleep. *Journal of Sleep Research*.

Emonson, D., & Vanderbeek, R. (1995). The use of amphetamines in U.S. Air Force tactical operations during desert shield and storm. *Aviation, Space, and Environmental Medicine*, **66**, 260-263.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, **98**, 506-528.

Hockey, G. (1986). Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. In K. Boff, R. Kaufman, and J. Thomas (Eds.), *Handbook of perception and human performance, Vol2*. (pp. 44-1-44-9). Westport.

Keren, G. (1991). Calibration and probability judgments: Conceptual and methodological issues. *Acta Psychologica*, **77**, 217-273.

Kreuger, G.P. (1989). Sustained work, fatigue, sleep loss and performance: a review of the issues. *Work & Stress*, **3**, 129-141.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Lyons, T., & French, J. (1991). Modafinil: The unique properties of a new stimulant. *Aviation, Space, and Environmental Medicine*, **62**, 432-435.

Pigeau, R., Naitoh, P., Buguet, A., McCann, C., Baranski, J., Taylor, M., Thompson. M. & Mack, I. (In Press). Modafinil, d-Amphetamine and placebo during 64 hours of sustained mental work. I: Effects on mood, fatigue, cognitive performance and body temperature. *Journal of Sleep Research*.

Stone, D. (1994). Overconfidence in initial self-efficacy judgments: Effects on decision processes and performance. *Organizational Behavior and Human Decision Processes*, **59**, 452-474.

Warot, D., Corruble, E., Payan, C., Weil, J. S., & Puech, A. J. (1993). Subjective effects of Modafinil, a new central adrenergic stimulant in healthy volunteers: a comparison with amphetamine, caffeine and placebo. *European Psychiatry*, **8**, 201-208.

#506692