Sleep
11(6):528–536, Raven Press, Ltd., New York

Sleep Loss and "Divergent" Thinking Ability

J. A. Horne

Human Sciences Department, Loughborough University, Leicestershire, England

Summary: Although much is known about the impact of sleep loss on many aspects of psychological performance, the effects on divergent ("creative") thinking has received little attention. Twelve subjects went 32 h without sleep, and 12 others acted as normally sleeping controls. All subjects were assessed on the figural and verbal versions of the Torrance Tests of Creative Thinking. As compared with the control condition, sleep loss impaired performance on all test scales (e.g., "flexibility," the ability to change strategy, and "originality," generation of unusual ideas) for both versions, even on an initial 5-min test component. In an attempt at further understanding of whether these findings might be explained solely by a loss of motivation, two additional short and stimulating tests were also used—a word fluency task incorporating high incentive to do well and a challenging nonverbal planning test. Performance at these tasks was still significantly impaired by sleep loss. Increased perseveration was clearly apparent. Apparently, 1 night of sleep loss can affect divergent thinking. This contrasts with the outcome for convergent thinking tasks, which are more resilient to short-term sleep loss. **Key Words:** Sleep loss—Divergent thinking—Creative thinking.

In sleep deprivation research, the emphasis on psychological performance measurement has been placed on the tests that are most sensitive to sleepiness, easy to score, and not prone to large practice effects that would cancel out any impairment caused by sleepiness. Therefore, the preferred tests are dull, monotonous, and unstimulating, such as 10–20 min of simple reaction time or vigilance testing (1–4). These tests are more inclined to assess the motivation to perform rather than the inherent capacity to do so. But because such tests readily show significant effects of sleep loss, we cannot assume that they are also measuring behaviours fundamental to the understanding of sleep function.

Complex tests such as reading comprehension, IQ performance, and decision making show little effect of total sleep loss until the second night (5–9). Such tasks are absorbing and stimulate one to apply more compensatory effort to counteract motivational decline. However, the tasks tap only certain aspects of performance. In providing a selection of possible answers, IQ tests involve convergent thinking, encourage the sub-

Accepted for publication April 1988.
Address correspondence and reprint requests to J. A. Horne at Human Sciences Department, Loughborough University, Leicestershire LE11 3TU, England.

ject progressively to "home in" on the solution, and rely on skills that are already well established. These tests are relatively easy to score.

Little is known about the effects of sleep loss on divergent thinking, i.e., skills requiring spontaneity, creativity, and flexibility. Tests of this type are tedious to score and have largely been ignored by sleep deprivation research. Twenty-five years ago, the Walter Reed studies (10) showed (e.g., using the "Cognitive Disorganisation" scale) that spontaneous thinking deteriorated rapidly during sleep deprivation. The scale is not precise and is now seldom used.

A more recently developed and sophisticated measure of divergent thinking is the Torrance Tests of Creative Thinking (11) (see below). These tests have not been used in sleep loss studies. Unfortunately, each set of tests takes 30–40 min to complete; within the context of sleep loss, this is a long time. Thus, any deterioration in creative thinking might result from falling subject interest and motivation rather than from decrements in inherent capacity for thinking in this manner. To overcome any motivational decline, one might address this issue by giving subjects incentives to do well. A remaining decrement might then be caused mostly by impaired capacity. The Torrance tests are complex, however, and incentives such as immediate knowledge of results and money for correct results, sometimes used in sleep loss studies for simpler performance tasks (12–14), are difficult to apply (e.g., which of the three scoring dimensions, described below, would be rewarded?).

The present study investigated whether divergent thinking was impaired by short-term sleep loss and attempted to gauge whether this might be caused by a motivational decline or by a real deficit in the subjects' capacity to think in this manner. However, a variety of factors affect the psychological performance of the sleep-deprived individual; these are clearly described elsewhere (3). Nevertheless, one of the most important factors is the subject's motivation. After 1 night of sleep loss, subjects can maintain normal levels of motivation and performance for ~5 min on the dullest of tests before a decline is apparent (12). With these points in mind, we selected three short-duration tests, designed to minimise the adverse effects of any decline in motivation, which tapped aspects of divergent thinking ability.

The first test consisted of the figural and verbal modes of the Torrance tests, which contain a challenging variety of subtests that encourage sleepy subjects to apply themselves throughout most of the entire test session. However, the first subtest of the verbal series (see below) lasts 5 minutes.

The second test was Shallice's nonverbal planning ("Tower of London") test (15), which is a short-duration task requiring flexibility in thinking style (see below) and can easily lead to perseveration in behaviour. The test is untried within the context of sleep loss, but it produces a very high level of interest and enthusiasm in the subject.

The third test adopted was Thurstone's word fluency test (see below). As many words as possible, beginning with a certain letter, are generated within a set time (16). The task is straightforward, and additional incentives can easily be added to encourage good performance.

METHOD

Tests

Torrance test of creative thinking. The Torrance test of creative thinking has two modes, verbal and figural, each having two matched forms (A and B). There are 30- and

45-min time limits for the figural and verbal modes, respectively. The verbal mode has seven subtests; the figural mode has three. Examples of these tests are as follows.

Verbal mode

1. A picture is presented of a person doing an apparently meaningless act. The respondent must: (a) write down all the questions that come to mind that might help explain the situation, (b) list as many possible causes of the event, and (c) list as many consequences of the event.
2. How many uses can a cardboard box be put to?
3. What would be the consequences of some hitherto unimaginable event—for example, what if clouds had string attached to them hanging down to earth?

Figural mode

1. Incomplete abstract figures are presented, and the respondent must complete these in a meaningful way.
2. Pages of small circles are presented and must be incorporated into as many different and simple pictures as possible.

For both verbal and figural modes, scores across each subtest are integrated into an overall test score covering three (verbal) or four (figural) dimensions: Flexibility, the ability to produce a variety of kinds of ideas, to shift from one approach to another, or to use a variety of strategies; Originality, the potential to produce ideas away from the obvious and commonplace; Elaboration, (used only in the figural form of the test) the ability to develop and embellish otherwise elaborate ideas; and Fluency, the number of ideas produced. Fluency helps with interpretation of the other three scores, as high fluency scores can be obtained at the expense of the other categories. For example, the banal thinker can score well on fluency but score low on everything else. The Torrance tests include detailed scoring manuals that provide high interrater reliability.

Nonverbal planning test. This apparatus consists of three pegs of different lengths and three beads of different colors. The beads must be moved in certain sequences peg to peg, from a standard configuration, to produce a set pattern within a specified number of moves. The test must be administered according to strict guidelines. Twelve presentations plus a demonstration example are possible, each having one solution varying between three and five moves. Strategies for the solutions change in the series and require a flexible and adaptable approach. In this respect, the test is a good guide to perseveration (difficulty in changing strategies). Two scores are obtained for each presentation: Planning time (studying the problem) from presentation to start of moving the beads, and solution time (physical action) from start of moving beads to the correct solution (most normal subjects tend not to begin this phase until they have the solution in mind).

For adult students, the combined times take ~90 s/presentation, and the entire test is completed within 15–20 min. The task was originally designed for single-measurement neuropsychological assessment. But because we wanted repeated measurement (see below), we split the test into halves, equated for difficulty and number of moves, i.e., 2 six-item subtests using presentations 1, 4, 6, 8, 9, and 12 ("form A") and the remaining presentations ("form B").

Thurstone word fluency test. Different letters (four for "any-length word", and four for "four-letter words") were selected for having similar word production rates (see below).

Anagrams. This test is a short verbal test of convergent thinking, used here simply

for comparison with the divergent thinking tasks. Words were selected from the "A.A." category of the Thorndike and Lorge (17) word list and were scrambled randomly. A pilot study (see below) showed that in 10 min ~7 six-letter words could be solved from a 10-item list. Two such lists were eventually produced. We emphasise that our study was not designed to compare divergent and convergent thinking tests, but concentrated instead on the former. The anagram test was stimulating, simply acting as a guide and reference point to the state of convergent thinking. The sleep deprivation literature strongly indicates that such a test would be insensitive to 1 night of sleep loss (1–4).

Pilot study and reliabilities of test scoring

A pilot study of 10 subjects under nonsleep deprivation conditions was run to develop some of the tasks, to check the test–retest reliabilities of the two forms of the nonverbal planning test and the A and B forms of both the figural and verbal modes of the Torrance test. Although Torrance (11) claims a high degree of reliability, practice/adaptation effects exist owing to the unusual nature of the tests. Half the subjects performed the forms of the planning and Torrance tests in the order A-B; the other subjects performed them in the order B-A, after an interval of half a day. For the nonverbal planning task, there was an improvement in the second testing session, similar for both A-B and B-A (mean improvement in planning time 28%, solution time 21%). For the Torrance tests, fluency and flexibility scores for both the verbal and figural dimensions improved by between 20 and 25% on the second setting and, respectively, by an average of 35 and 50% for figural and verbal originality. For both figural and verbal modes, no difference was discernible in the degree of improvement whether form A preceded or followed form B. Obviously, large adaptation effects occur, and the validity and applicability of the Torrance test for the present study are questionable. However, despite these limitations, the Torrance test is probably the best known and most comprehensive test of creativity available in the United Kingdom, and one must accept its limitations. Obviously, the design of the main study had to try to control for these adaptation effects. The pilot study was also used to try out lists of anagrams and to check the initial letters used for the Thurstone word fluency task. The selection of letters available for the four-letter words were G, L, W, and H; for any-length words, the selection was M, A, B, and D (further details are given below).

Main study

Subjects. Twenty-four subjects (15 women and 9 men aged 19–23 years) were recruited by advertisement from the campus student population. All were screened for being physically and mentally healthy, medication-free, good sleepers (7.5–8.0 h), and nonsmokers, of average fitness. They were paid for their services. They were randomly divided into two equal groups: sleep-deprived and control. Subjects were carefully briefed.

Design and procedure. Four experimental runs were performed at weekly intervals on groups of six subjects, beginning late Friday afternoons. Each run contained a new group of subjects. The first group comprised six subjects deprived of sleep that night; the second comprised a control group allowed to sleep normally, followed by another control group, and finally the remaining deprivation group. Subjects were familiarised with the test material on the Thursday before the test. For the four previous nights, subjects had to go to bed no later than 2330 h. They kept sleep logs. Runs began at 1630 h on the test day (Friday) and ended at 2100 h the next day. Except for the sleep

period in the control group, all subjects remained within the vicinity of the laboratory and were under constant supervision throughout the study. They ate in the laboratory at 1830, 0800, and 1300 h and had an additional snack at 0200 h during sleep deprivation. Alcohol, coffee, and cola drinks were banned. Fruit juice and weak tea was freely available. Control subjects went to their campus residences at 2300 h on the test day and had to retire by 2330 h. They returned to the laboratory at 0800 h the next day. Each hour during wakefulness, all subjects rated themselves on the Stanford Sleepiness Scale (18).

The schedule of testing for Friday and Saturday is shown in Fig. 1. Between 1630 h and 1800 h on Friday, subjects were individually assessed on forms A or B of the nonverbal planning test. This was repeated at the same time the next day, using the other form. Half the subjects within each group underwent A–B; the others underwent B–A. Between 1900 and 2100 h on both days, subjects were given (in group sessions) forms A or B of each of the verbal and figural versions of the Torrance tests. There was a 30-min interim break. These tests were given in a balanced order. The instructions given in the manuals were rigidly adhered to; scoring was done on a blind basis by two scorers following the detailed procedure given in the manuals, and the two assessments were averaged for each result. The anagrams test, which lasted 10 min, was given in group sessions on two occasions, at 1815 h on both Friday and Saturday.

Two forms of Thurstone's word fluency test were presented: generation of words of (a) any length, (b) four letters. The tests were consecutive, used different letters, and each lasted 5 min. Four letters were available for both (a) and (b) (see above). For each subtest, however, every subject was randomly assigned only two of the possible four letters, one of these being for Friday and the other for Saturday. The Thurstone tests were administered at 1800 h on Friday and Saturday. The Friday session had no reward. Just before the Saturday session, sleep-deprived and control subjects were told that they would receive 2 pence (~4 cents) for every correct word generated, and that £1.5 (~$3) could be earned in 10 min (Note: The cost of living is lower in the United Kingdom than in the United States; e.g., an average wage for a student doing parttime work is ~$4/h). A competition atmosphere was allowed to prevail among each group for this test. The tests were scored immediately, and payment was given on the spot. No knowledge of results or other form of feedback was given for the other tests. The rest of Saturday morning and afternoon was taken up mostly in leisure pursuits.

Analysis of data. The two groups of subjects were from the same pool and were randomly assigned to their respective groups. Data from the nonverbal planning, Torrance, anagram, and Thurstone tests were each analysed by the following method. Using the control group data only, a regression equation was calculated for the findings for Friday versus those for Saturday. This represented the predictable change owing to practice, etc. or, in the case of the Thurstone scores, practice plus reward. The equa-

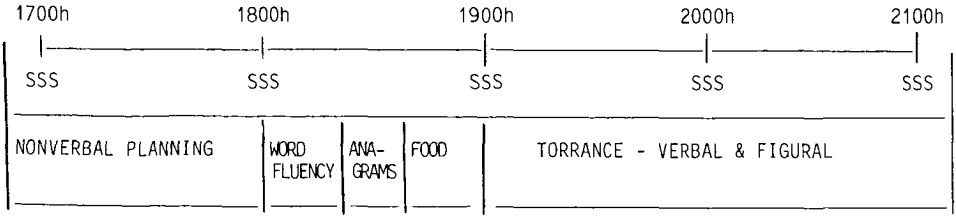

FIG. 1. Schedule of testing during the evening for Friday and Saturday. Sleepiness ratings on the Stanford Sleepiness Scale (SSS) were performed hourly throughout wakefulness.

tion was then applied to the respective Friday (baseline) data from the sleep-deprived group to give a predicted value for each subject's score. These results were compared with the actual Saturday (sleep loss) scores, using a one-way analysis of variance (ANOVA). Analyses of covariance were not performed on the data from the control and sleep-deprived subjects combined because in most cases the regressions from the two groups were nonhomogeneous since sleep loss usually counteracted practice effects.

RESULTS

The findings for all performance tests are given in Table 1, and significant values (two-tail) are identified according to the criteria given above. Trends for sleepiness ratings are shown in Fig. 2. The expected increases in sleepiness following sleep loss are clearly evident.

For all tests, baselines of the two groups showed notable similarities (Table 1). The nonverbal planning test produced significant increases in planning time following sleep

TABLE 1. *Findings with performance tasks, with significance levels*

		Day 1 (baseline)		Day 2	
Test		Control	Sleep deprived	Control	Sleep deprived
Nonverbal planning					
Planning time (s)	x	60.0	61.3	46.9	89.7[a]
	σ	7.7	7.9	7.4	18.6
Solution time (s)	x	42.8	46.5	36.5	54.6
	σ	9.0	10.1	6.8	11.4
Torrance—figural					
Fluency (scale rating)	x	19.7	20.9	24.5	15.7[b]
	σ	8.6	3.7	8.9	5.0
Flexibility (scale rating)	x	16.0	17.3	19.7	13.0[c]
	σ	4.7	2.9	4.1	3.2
Originality (scale rating)	x	32.8	28.7	46.5	16.2[c]
	σ	9.2	6.2	7.6	4.1
Elaboration (scale rating)	x	70.0	86.0	88.4	57.7[c]
	σ	22.7	26.6	18.7	17.8
Torrance—verbal					
Fluency (scale rating)	x	79.7	85.2	95.3	62.3[c]
	σ	19.4	19.8	16.6	9.8
Flexibility (scale rating)	x	37.2	39.8	45.5	26.6[c]
	σ	8.7	9.4	5.8	4.9
Originality (scale rating)	x	66.3	71.5	84.5	38.5[c]
	σ	16.6	16.3	19.3	8.4
Thurstone word fluency					
Any-length words (no. of words)	x	39.3	38.2	52.4	40.2[a]
	σ	5.8	3.5	9.7	3.6
Four-letter words (no. of words)	x	27.7	24.5	25.0	21.5
	σ	4.5	7.1	7.4	2.9
Anagrams					
No. solved	x	7.6	6.8	8.3	7.1
	σ	2.1	1.5	1.7	2.3

Findings are means (x) and SD (σ). Details given in text.

Significant difference between sleep-deprived and control groups on day 2: [a]$p < 0.05$, [b]$p < 0.02$, [c]$p < 0.002$.

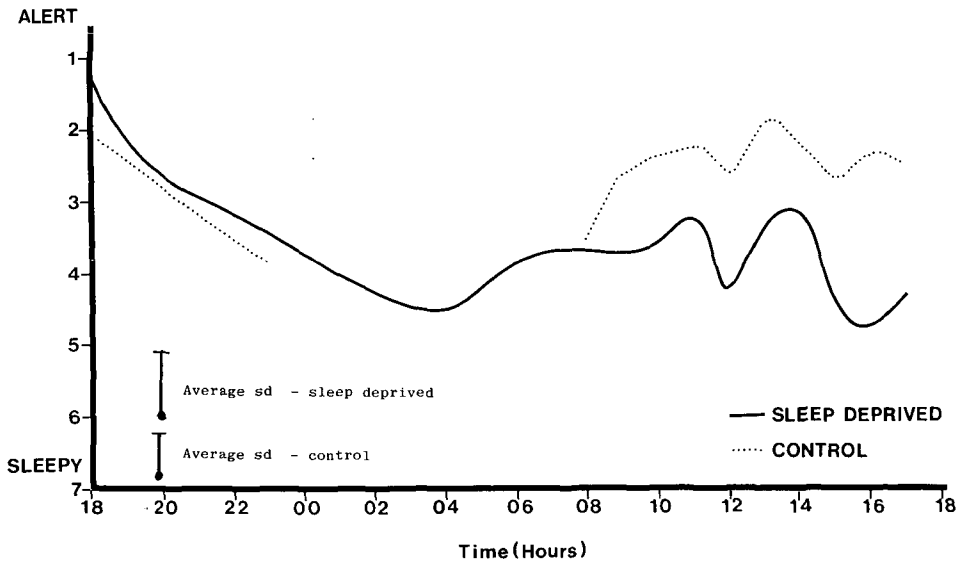

FIG. 2. Group hourly means of self-ratings of alertness on the Stanford Sleepiness Scale, with average SD.

loss ($F = 8.24$, df 1,11: $p < 0.05$), where the two means differed by a magnitude of 2. Solution time was not significantly affected. Sleep loss led to significant falls in all four scales in the Torrance Figural Test: Elaboration ($F = 23.1$, df 1,11: $p < 0.002$), Originality ($F = 28.3$, df 1,11: $p < 0.002$), Flexibility ($F = 20.5$, df 1,11: $p < 0.002$), and Fluency ($F = 13.8$, df 1,11: $p < 0.02$). All three scales in the Torrance Verbal Test were adversely affected by sleep loss: Originality ($F = 38.8$, df 1,11: $p < 0.002$), Flexibility ($F = 29.0$, df 1,11: $p < 0.002$), and Fluency ($F = 20.1$, df 1,11: $p < 0.002$). Although test 1 (described above) of the Torrance Verbal series (Table 2) showed an improvement from day 1 to day 2 for both groups, improvement was significantly greater for the control group and for all three dimensions: Fluency ($F = 13.8$, df 1,11: $p < 0.02$), Flexibility ($F = 7.84$, df 1,11: $p < 0.05$), and Originality ($F = 12.4$, df 1,11: $p < 0.02$).

The word fluency task produced no significant findings for four-letter words. However, for words of any length, there was a significant outcome ($F = 8.9$, df 1,11: $p < 0.05$), with reward failing to produce a rise in words generated by the sleep-deprived group. Anagrams gave no significant results.

DISCUSSION

Thirty-two hours without sleep had a substantial impact on most aspects of divergent thinking measured. All scales of both the verbal and figural forms of the Torrance test

TABLE 2. *Means and SD of changes from day 1 to day 2 in scores for task 1 on Torrance verbal scale*

Test	Control	Sleep deprived
Fluency	+ 47.4% (σ 21.1%)	+ 21.2% (σ 12.5%)[a]
Flexibility	+ 52.2% (σ 17.1%)	+ 31.8% (σ 9.5%)[b]
Originality	+ 52.0% (σ 15.0%)	+ 19.0% (σ 16.2%)[a]

Significant difference between sleep-deprived and control groups on day 2: [a]$p < 0.02$, [b]$p < 0.05$.

were substantially influenced by sleep loss. However, a large masking by a practice/adaption effect occurred, whereby all scores were liable to improve at the second sitting. This was clearly apparent in the pilot study and in the control group of the main study. Although subjects were shown the principles of the Torrance test and practiced at a dummy version (constructed by us) during a preliminary familiarisation day, this was not adequate to overcome subsequent practice effects.

Only two forms of each version of the Torrance tests are available; we reserved them for days 1 and 2 of the main study. Even if there had been more formal versions of the test, practice effects probably would have continued because the test is novel and relatively complex. Consequently, prior to a main study, the typical procedure of repeated practice at a performance task to allow the subjects to asymptote beforehand (the simpler the test, the more feasible this becomes) could not have been followed. Instead, we tried to estimate the practice effect using a control group.

Unwanted motivation effects should have been at their lowest during the first 5-min subtest of the verbal form (Table 2). Although the sleep-deprived group still showed some improvements after sleep loss, in the subtest the improvements were much lower than those of the control group. This suggests that sleep loss was affecting mechanisms other than those concerned with motivation and subject interest. Such a conclusion is supported by the findings from the short-duration and stimulating nonverbal planning test. Planning time showed by far the greater effect and is the period when most subjects solve or nearly solve the problem in their mind, whereas solution time is the period of putting problem-solving into practice. From our observations and the subjects' own reports, sleep loss made them fixate on previously successful strategies when attempting solutions to the next problem (which generally required a new approach); i.e., they seemed to perseverate unduly.

Perseveration was also apparent in the Torrance and Thurstone tasks. For example, in the latter, it was common for sleep-deprived subjects to write down the same word repeatedly, and then to cross it out. This phenomenon of perseveration may well be a topic worthy of fuller investigation within the context of sleep loss (9). The nonverbal planning and Thurstone tasks are also used as neuropsychological tests of frontal lobe function (15,16,19).

In a recent study, Webb (4) reported that 2 nights of sleep loss did not affect an "object usage" test, which was very much like one of the Torrance tests we used. Webb's subjects were given only 2-min periods in which to describe all possible ways any given object could be used. Our own findings indicate that 2 min is too short a period for completion and may explain the different outcomes between the study of Webb and our study. At least 1 min is required for subjects to "collect their thoughts"; a few more minutes are needed for writing or verbalizing the ideas.

With regard to the present findings for word fluency, the reward had no obvious effect on the generation of four-letter words by either group. However, the control group showed a marked improvement for any-length words. The four-letter condition was apparently too constraining, whereas the any-length word condition was more open-ended and perhaps more "divergent." Nevertheless, despite the attractive incentive, the sleep-deprived subjects could not be induced to generate more words for the open-ended form.

Divergent thinking tests can incorporate tasks that may appear to be facile and pointless to certain subjects—even embarrassing to them; e.g., asking military personnel or someone who thinks in an "unimaginative" way to respond to some of the

Torrance subtests (see Method section) might result in inhibited responses under both baseline and sleep deprivation conditions. Consequently, sleep loss effects might be masked. Our subjects (partly perhaps because they were students), appeared to enter into the spirit of these tests and, as far as we could determine, were not inhibited from doing their best. We stress that it is important to explain these tests fully to the subjects beforehand and to obtain their full confidence. Good tests of divergent thinking ability are few and, for military personnel for example, divergent thinking tests that have a more realistic setting must be devised.

The limited findings concerning the single test of convergent thinking we used tend to confirm the outcome described in the literature (1–4,8,9). Such a cognitive style is not affected by short-term sleep loss. On the other hand, other results from the present study appear to point to impaired divergent and flexible thinking. This seems only partly attributable to motivational changes (e.g., a decline in the subjects' interest in the tasks), as short-duration stimulating tests were also affected despite the subjects' best efforts to perform well. Something more fundamental apparently occurs with this form of behaviour during sleep loss.

Acknowledgment: I thank Julie Todd for help with data collection and analysis.

REFERENCES

1. Wilkinson RT. Sleep deprivation. In: Edholm OG, Bacharach A, eds. *The physiology of human survival.* New York: Academic Press, 1965:399–430.
2. Naitoh P. Sleep deprivation in human subjects: a reappraisal. *Waking Sleeping* 1976;1:56–60.
3. Johnson LC. Sleep deprivation and performance. In: Webb WB, ed. *Biological rhythms, sleep, and performance.* New York: Wiley, 1982:111–41.
4. Webb WB. Experiments on extended performance: repetition, age, and limited sleep periods. *Behav Res Methods, Instruments Computers* 1985;17:27–36.
5. Edwards AS. The effect of the loss of one hundred hours of sleep. *Am J Psychol* 1941;54:80–91.
6. Carmichael L, Kennedy JL, Mead LC. Some recent approaches to the experimental study of human fatigue. *Proc Natl Acad Sci USA* 1949;35:691–6.
7. Percival JE, Horne JA, Tilley AJ. Effects of sleep deprivation on tests of higher cerebral functioning. In: Koella WP, ed. *Sleep 1982.* Basel: Karger, 1983:390–1.
8. Webb WB. Sleep deprivation and reading comprehension. *Biol Psychol* 1986;22:169–72.
9. Horne JA. *Why we sleep: the functions of sleep in humans and other mammals.* Oxford: Oxford University Press, 1988.
10. Morris GO, Williams HL, Lubin A. Misperception and disorientation during sleep deprivation. *Arch Gen Psychiatry* 1960;2:247–54.
11. Torrance EP. *Torrance tests of creative thinking—norms and technical manual.* Lexington, Massachusetts: Ginn, 1974.
12. Williams HL, Lubin A, Goodnow JJ. Impaired performance with acute sleep loss. *Psychol Monographs* 1959;73:(whole no. 484).
13. Wilkinson RT. Interaction of sleep loss with knowledge of results, repeated testing and individual differences. *J Exp Psychol* 1961;66:332–7.
14. Horne JA, Pettitt AN. High incentive effects on vigilance performance during 72 hours of total sleep deprivation. *Acta Psychol* 1985;58:133–9.
15. Shallice T. Specific impairments of planning. *Philos Trans R Soc London B* 1982;298:199–209.
16. Pendelton M, Heaton R, Lehman R, Hulihan D. Diagnostic utility of the Thurstone Word Fluency Test in neuropsychological evaluations. *J Clin Neuropsychol* 1982;4:307–17.
17. Thorndike E, Lorge I. *The teachers word book of 30,000 words.* New York: Columbia University Press, 1944.
18. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. *Psychophysiology* 1973;10:431–6.
19. Kolb B, Whishaw IQ. *Fundamentals of human neuropsychology,* 2nd ed. San Francisco: W. H. Freeman, 1985.